

Wed Aug 20 13:35:30 2003

us-09-512-082-19.rpr

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: August 20, 2003, 12:25:56 ; Search time 81.0602 seconds
(without alignments)
(137.621 Million cell updates/sec)

Title: US-09-512-082-19
Perfect score: 608
Sequence: 1 EVOLLESGGGLVPGGSLRL.....AKPFYEDYWGGLTVTVSS 116

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616862 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR76:*

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556.5	91.5	140	S31686	Ig heavy chain V r
2	552.5	90.9	140	S31588	Ig heavy chain - h
3	551.5	90.7	119	S31107	Ig heavy chain - h
4	551.5	90.7	123	S31114	Ig heavy chain V r
5	549.5	90.4	119	C36005	Ig heavy chain - h
6	545.5	89.7	138	S31108	Ig heavy chain V r
7	545.5	89.6	120	S31666	Ig heavy chain V r
8	544	89.5	124	S48798	Ig heavy chain V r
9	544	89.5	119	S20782	Ig heavy chain V r
10	538.5	88.6	121	D36005	Ig heavy chain - h
11	535.5	88.1	127	T55673	Ig heavy chain V r
12	535.5	88.1	127	S31689	Ig heavy chain pre
13	526.5	86.6	134	S31699	Ig heavy chain V r
14	526.5	86.6	160	S05271	Ig heavy chain V r
15	524.5	85.5	125	S30531	Ig heavy chain V r
16	520	85.5	112	S21624	Ig heavy chain V r
17	512.5	84.3	143	S31669	Ig heavy chain V r
18	509	83.7	141	S31113	Ig heavy chain - h
19	508.5	83.6	121	S26794	Ig heavy chain V r
20	507.5	83.3	123	S31603	Ig heavy chain V r
21	506.5	82.8	132	S31603	Ig heavy chain pre
22	503.5	82.8	140	A30532	Ig heavy chain V r
23	502	82.6	114	S31120	Ig heavy chain V r
24	501	82.4	128	S26790	Ig heavy chain V r
25	500	82.2	108	PH1648	Ig heavy chain V-D
26	500	82.0	120	S44111	Ig heavy chain V r
27	500	82.0	109	PH1649	Ig heavy chain V r
28	498.5	82.0	121	S19666	Ig heavy chain V r
29	498.5	82.0	121	S19666	Ig heavy chain V r

30	498	81.9	140	2	S70442	Ig heavy chain pre
31	497	81.7	122	2	S20772	Ig heavy chain V r
32	495.5	81.5	135	2	S31598	Ig heavy chain V r
33	492.5	81.0	121	2	G36005	Ig heavy chain V r
34	492.5	81.0	123	2	S30532	Ig heavy chain - h
35	492	80.9	122	2	S31117	Ig heavy chain V r
36	491	80.8	114	2	S46390	Ig heavy chain (su
37	490	80.6	118	2	S31105	Ig variable region
38	489	80.6	147	2	S31780	Ig heavy chain V r
39	489	80.4	122	2	S31675	Ig heavy chain V r
40	487	80.1	98	2	S26889	Ig heavy chain pre
41	487	80.0	117	2	A45953	Ig heavy chain V r
42	486.5	80.0	120	2	S36273	Ig heavy chain V r
43	486	79.9	128	2	S46391	Ig heavy chain V r
44	486	79.9	128	2	S31595	Ig heavy chain pre
45	486	79.9	151	2	A60943	Ig heavy chain pre

ALIGNMENTS

RESULT 1

S31686 Ig heavy chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S31686
R:Guthrie, A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelle, C.
Submitted to the EMBL Data Library, June 1992

A:Description: Mechanisms that generate human immunoglobulin diversity operate from t.
A:Reference number: S31585

A:Accession: S31686
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-140 <CU>
A:Cross-references: EMBL:Z14205; NID:930969; PIDN:CA78574.1; PID:930970

C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 91.5%; Score 556.5; DB 2; Length 140;
Best Local Similarity 90.1%; Pred. No. 6.7e-43;
Matches 109; Conservative 5; Mismatches 2; Indels 5; Gaps 2;

QY	1	EVOLLESGGGLVPGGSLRLCAAGFTFFSSFSMSVNRQAPGKLEWVSISGSGSTYY	60
Db	20	EVOLLESGGGLVPGGSLRLCAAGFTFFSSFSMSVNRQAPGKLEWVSISGSGSTYY	79
QY	61	ADSVKGRFTISRDNSKNTLYLQNMNSLRADTAAYTCAR-PF-----PYPDYWGGLTVTVSS	115
Db	80	SDSVKGRFTISRDNSKNTLYLQNMNSLRADTAAYTCARCFAGSGSPFDYWGGLTVTVSS	139
QY	116	S 116	
Db	140	S 140	

RESULT 2
S31588 Ig heavy chain V region - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

R:Guthrie, A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelle, C.
Submitted to the EMBL Data Library, June 1992

A:Description: Mechanisms that generate human immunoglobulin diversity operate fro

A:Reference number: S31585

A:Accession: S31588
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-140 <CU>
A:Cross-references: EMBL:Z14200; NID:930957; PIDN:CA78569.1; PID:930958

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 90.9%; Score 552.5; DB 2; Length 140;
Matches 107; Conservative 6; Mismatches 3; Indels 5; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
DB 20 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
QY 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 79
DB 80 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 115
QY 116 S 116
DB 140 S 140

RESULT 3

S31107

Ig heavy chain - human

C:Species: Homo sapiens (man)

C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999

R:Accession: S31107

R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurmat

A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple

A:Reference number: S31104; MUID:92111633; PMID:1730252

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-119 <RAA>

A:Cross-references: EMBL:X62955

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 90.7%; Score 551.5; DB 2; Length 119;
Matches 108; Conservative 5; Mismatches 2; Indels 5; Gaps 2;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
DB 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
QY 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 116
DB 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 119

RESULT 4

S31114

Ig heavy chain - human

C:Species: Homo sapiens (man)

C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999

R:Accession: S31114

R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurmat

A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple

A:Reference number: S31104; MUID:92111633; PMID:1730252

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-123 <RAA>

A:Cross-references: EMBL:X62963

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 90.7%; Score 551.5; DB 2; Length 123;
Matches 108; Conservative 4; Mismatches 4; Indels 7; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
DB 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
QY 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 79
DB 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 113
QY 114 VSS 116
DB 121 VSS 123

RESULT 5

C36005

Ig heavy chain V region (30p1) - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1990 #sequence_revision 21-Dec-1990 #text_change 16-Aug-1996

R:Accession: C36005

R:Schröder, Jr., H.W.; Wang, J.Y.

A:Title: Preferential utilization of conserved immunoglobulin heavy chain variable ge

A:Reference number: A36005; MUID:9034571; PMID:2117273

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-119 <SCB>

A:Cross-references: GB:M18513

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 90.4%; Score 549.5; DB 2; Length 119;
Matches 107; Conservative 5; Mismatches 4; Indels 3; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
DB 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVROAPGKLEWVSISGSGTTY 60
QY 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 116
DB 61 ADSVKGRTISRDNSKNTLYLQMSLRADTAAYVYCAKPPY-----FDYWGQGLTVTVSS 119

RESULT 6

S31108

Ig heavy chain - human

C:Species: Homo sapiens (man)

C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999

R:Accession: S31108

R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurmat

A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple

A:Reference number: S31104; MUID:92111633; PMID:1730252

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-119 <RAA>

A:Cross-references: EMBL:X62956

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 89.7%; Score 545.5; DB 2; Length 119;
Matches 107; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

Oy 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVROAPGKLEWSSISGSGTTY 60
1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Db 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Oy 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRP---FPYFDWGGGTLTVSS 116
61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDRRLGTGTFYMGGGLTVYSS 119
Db 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDRRLGTGTFYMGGGLTVYSS 119

RESULT 7

31666
Ig heavy chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31666
R:Curator: A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelie, C.
Submitted to the EMBL Data Library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
A:Reference number: S31585
A:Accession: S31666
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-138 <CUI>
A:Cross-references: EMBL:Z14202; NID:g30963; PIDN:CAA78571.1; PID:g30964
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 89.7%; Score 545.5; DB 2; Length 138;
Best Local Similarity 89.1%; Pred. No. 6.4e-42;

Matches 106; Conservative 6; Mismatches 4; Indels 3; Gaps 1;

Oy 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVROAPGKLEWSSISGSGTTY 60
1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 79
Db 20 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 79
Oy 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRP---FPYFDWGGGTLTVSS 116
61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRTGYWFDLMRGRTLVYSS 138
Db 80 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRTGYWFDLMRGRTLVYSS 138

RESULT 8

S48798
Ig heavy chain V region (anti-Sm, VH3/DxP4/JHb) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S48798
R:Mahmoudi, M.; Edwards, J.; Cairns, E.; Bell, D.
Submitted to the EMBL Data Library, October 1994
A:Description: Molecular characterization of natural human anti-Sm autoantibodies.
A:Reference number: S48797
A:Accession: S48798
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <MAN>
A:Cross-references: EMBL:Z46382; NID:g562324; PIDN:CAA86521.1; PID:g1340167
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 89.6%; Score 545; DB 2; Length 120;
Best Local Similarity 89.2%; Pred. No. 6.1e-42;
Matches 107; Conservative 4; Mismatches 5; Indels 4; Gaps 1;

Oy 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVROAPGKLEWSSISGSGTTY 60
1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Db 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Oy 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRP---FPYFDWGGGTLTVSS 116
61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDRGFMGSGYDYGCGGTLTVYSS 120
Db 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDRGFMGSGYDYGCGGTLTVYSS 120

RESULT 9

S20782
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C:Accession: S20782
R:Mortari, F.; Wang, J.; Schroeder, H.W.
Submitted to the EMBL Data Library, April 1992
A:Description: Analysis of the IgA and IgG rearranged VH repertoire of human cord blo
A:Reference number: S20765
A:Accession: S20782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <MOR>
A:Cross-references: EMBL:Z11946; NID:g33897; PIDN:CAA78003.1; PID:g33898
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 89.5%; Score 544; DB 2; Length 124;
Best Local Similarity 84.7%; Pred. No. 7.8e-42;
Matches 105; Conservative 7; Mismatches 4; Indels 8; Gaps 1;

Oy 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVROAPGKLEWSSISGSGTTY 60
1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSISGSGSTYY 60
Db 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSISGSGSTYY 60
Oy 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRP---FPYFDWGGGTLTV 112
61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKERRIAFGVYIPHEWYGGGTLTV 120
Db 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKERRIAFGVYIPHEWYGGGTLTV 120
Oy 113 TVSS 116
111
Db 121 TVSS 124

RESULT 10

D36005
Ig heavy chain V region (M43) - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1990 #sequence_revision 21-Dec-1990 #text_change 16-Dec-1998
C:Accession: D36005
R:Schroeder Jr., H.W.; Wang, J.Y.
Proc. Natl. Acad. Sci. U.S.A. 87, 6146-6150, 1990
A:Title: Preferential utilization of conserved immunoglobulin heavy chain variable g
A:Reference number: A36005; MUID:90349571; PMID:2117273
A:Accession: D36005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-119 <SCH>
A:Cross-references: GB:M34024
A:Gene: IGHG; IGHV1
A:Map position: 14q32.33-14q32.33
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 88.6%; Score 538.5; DB 2; Length 119;
Best Local Similarity 89.1%; Pred. No. 2.3e-41;
Matches 106; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

Oy 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVROAPGKLEWSSISGSGTTY 60
1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Db 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVROAPGKLEWSSAISGSGSTYY 60
Oy 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKRP---FPYFDWGGGTLTVSS 116
61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDMNDNFDWGGGTLTVYSS 119
Db 61 ADVKGRFTISRDNKNTLYLQNMNSLRADTAAYVYCAKDMNDNFDWGGGTLTVYSS 119

RESULT 11

155673
Ig heavy chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C:Accession: 155673
R:Kight, G.B.; Agnello, V.; Bonagura, V.; Barnes, J.L.; Panka, D.J.; Zhang, Q.X.
J. Exp. Med. 178, 1903-1911, 1993
A:Title: Human rheumatoid factor cross-idiotypes. IV. Studies on WA xid-positive Igm with
clint from the 17.109 and G6 Xids.
A:Reference number: 155673; MUID:94065558; PMID:8245772
A:Accession: 155673
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-121 <RES>
A:Cross-references: GB:M87268; NID:q186197; PIDN:AAC37536.1; PID:q186198
C:Gene: GDB:IGHM
A:Cross-references: GDB:120086; OMIM:147020
A:Map position: 14q32.33-14q32.33
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 88.1%; Score 535.5; DB 2; Length 121;
Best Local Similarity 86.0%; Pred. No. 4.4e-41;
Matches 104; Conservative 5; Mismatches 7; Indels 5; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
DB 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
QY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 115
DB 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 120

QY 116 S 116
DB 121 S 121

RESULT 12

338489
Ig heavy chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S38489
R:Marks, J.D.; Onehand, W.H.; Bye, J.M.; Finnern, R.; Gorlick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a F
A:Reference number: S38488
A:Accession: S38489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <MAR>
A:Cross-references: EMBL:Z23028; NID:9414025; PIDN:CAA80563.1; PID:9414026
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 88.1%; Score 535.5; DB 2; Length 127;
Best Local Similarity 81.9%; Pred. No. 4.6e-41;
Matches 104; Conservative 8; Mismatches 2; Indels 13; Gaps 2;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
DB 1 QVOLVSGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
QY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 107
DB 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 120
QY 108 QGLTVV 114
DB 121 QGLTVV 127

RESULT 13

331699
Ig heavy chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S31699
R:Cusinsler, A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelle, C.
submitted to the EMBL Data Library, June 1992
A:Description: Mechanisms that generate human immunoglobulin diversity operate from t
A:Reference number: S31695
A:Accession: S31699
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-134 <CU>
A:Cross-references: EMBL:Z14201; NID:930961; PIDN:CAA87570.1; PID:930962
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 86.6%; Score 526.5; DB 2; Length 134;
Best Local Similarity 87.9%; Pred. No. 3.1e-40;
Matches 102; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
DB 20 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 79
QY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 116
DB 80 SDVKGRLTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 134

RESULT 14

505271
Ig heavy chain precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Aug-1996
C:Accession: S05271; S04602
R:Kishimoto, T.
submitted to the EMBL Data Library, March 1989
A:Reference number: S05270
A:Accession: S05271
A:Molecule type: mRNA
A:Residues: 1-160 <RTS1>
A:Cross-references: EMBL:X14584
R:Kishimoto, T.; Okajima, H.; Okumoto, T.; Taniguchi, M.
Nucleic Acids Res. 17, 4385, 1989
A:Title: Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains
A:Reference number: S04601; MUID:89296497; PMID:2500644
A:Accession: S04602
A:Molecule type: mRNA
A:Residues: 1-144 <RTS2>
A:Cross-references: EMBL:X14584
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-160/Product: Ig heavy chain (fragment) #status predicted <MAT>
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 86.6%; Score 526.5; DB 2; Length 160;
Best Local Similarity 83.2%; Pred. No. 3.7e-40;
Matches 104; Conservative 7; Mismatches 5; Indels 9; Gaps 2;

QY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
DB 20 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 79
QY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 111
DB 80 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYCAKPPFDYWGQGLTVVSS 139

QY	112	VTVSS	116
Db	140	VTVSS	144

RESULT 15

ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Aug-1996
A:Accession: S30531
R:Marlette, X.
submitted to the EMBL Data Library, October 1992
A:Reference number: S30520
A:Accession: S30531
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-125 <MAR>
A:Cross-references: EMBL:Z18317
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
E:P15-98/Domain: Immunoglobulin homology <IMM>

Query Match	86.3%	Score 524.5	DB 2:	Length 125:
Best Local Similarity	81.6%	Pred. No. 4,3e-40:		
Matches 102;	Conservative	7;	Mismatches 7;	Indels 9; Gaps 1.
Oy	1	EYQLLESGGGLVQPGGSLRLSCAASGFFTFSSFSMSMVRQAPCGKLEWVSSISGSGSTYY	60	
		: : :		
Dd	1	EYQLVESGGGLVQPGGSLRLSCAASGFFTFSSYSMMWRQAPCKGLEIMISYISSSTIYY	60	
Oy	61	ADSVKGRFTISDNKSNTLYLQMSIRAEFDIVYICAKP-----PPRYDYGQGTLL	111	
		: : :		
Dd	61	ADSVKGRFTISRDKNAKNLSLYLMQNSLRADETLVYCYCARSRNYDSGSYSHYEDYQGQGTLL	120	
Oy	112	VTVSS	116	
Dd	121	VTVSS	125	

Search completed: August 20, 2003, 12:42:13
Job time : 82.0602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 80.3614 Seconds
(without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-19
Perfect score: 608
Sequence: 1 EVQVLESGGGLVQPGGSLRL.....AKFPYFDYGGGTLVTWSS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/ptodate/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodate/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodate/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodate/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodate/1/1aa/PCNUS.COMB.pep.*
6: /cgn2_6/ptodate/1/1aa/backfilled1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
And is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	544	89.5	131	3	US-08-983-607-28
2	539	88.7	120	4	US-09-025-769B-38
3	539	88.7	120	4	US-09-025-769B-63
4	539	88.7	281	4	US-09-025-769B-178
5	537.5	88.4	113	3	US-08-974-899-6
6	537.5	88.4	125	2	US-08-428-197-1
7	537.5	88.4	116	3	PCT-US93-10555-1
8	535	88.0	115	3	US-08-983-607-36
9	534.5	87.9	119	1	US-07-988-925-11
10	534.5	87.9	119	2	US-08-362-780-11
11	531	87.3	263	4	US-09-069-821-3
12	531	87.3	283	4	US-09-420-592A-6
13	529.5	87.1	249	4	US-10-039-785-53
14	528	86.8	122	2	US-07-934-373C-21
15	528	86.8	122	3	US-08-437-642B-21
16	528	86.8	122	4	US-08-146-206C-21
17	528	86.8	122	5	PCT-US93-07832-21
18	526	86.5	135	3	US-08-619-491-8
19	526	86.5	135	3	US-08-579-378A-20
20	525	86.3	135	3	PCT-US96-13152-4
21	522	85.9	135	5	PCT-US95-07302-8
22	521.5	85.8	125	1	US-08-478-039-99
23	521.5	85.8	125	1	US-08-476-349A-99
24	519.5	85.4	120	2	US-08-428-197-22
25	519.5	85.4	120	5	PCT-US93-10555-22
26	519	85.4	140	3	US-08-983-607-32
27	518.5	85.3	120	2	US-08-428-197-20

28	518.5	85.3	120	5	PCT-US93-10555-20	Sequence 20, Appl
29	515.5	84.8	117	3	US-08-983-607-46	Sequence 46, Appl
30	515.5	84.8	120	2	US-08-428-197-24	Sequence 26, Appl
31	515.5	84.8	120	2	US-08-428-197-26	Sequence 26, Appl
32	515.5	84.8	120	2	US-08-428-197-28	Sequence 28, Appl
33	515.5	84.8	120	5	PCT-US93-10555-24	Sequence 24, Appl
34	515.5	84.8	120	5	PCT-US93-10555-26	Sequence 26, Appl
35	515.5	84.8	120	5	PCT-US93-10555-28	Sequence 28, Appl
36	514	84.5	116	2	US-08-428-197-2	Sequence 2, Appl
37	514	84.5	116	3	PCT-US93-10555-2	Sequence 2, Appl
38	512	84.2	124	3	US-08-983-607-51	Sequence 31, Appl
39	511	84.0	122	5	PCT-US93-08435-12	Sequence 12, Appl
40	508.5	83.6	117	4	US-09-025-769B-24	Sequence 24, Appl
41	508	83.6	120	2	US-07-934-373C-4	Sequence 4, Appl
42	508	83.6	120	3	US-08-437-642B-4	Sequence 4, Appl
43	508	83.6	120	4	US-08-146-206C-4	Sequence 14, Appl
44	506	83.2	122	5	PCT-US93-08435-14	Sequence 14, Appl
45	506	83.2	144	1	US-08-026-320A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-983-607-28
Sequence 28, Application US/08983607
Patent No. 6140470
GENERAL INFORMATION:
APPLICANT: Xiahong Cai
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
TITLE OF INVENTION: bodies
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSER: Department of Molecular Biophysics
ADDRESSER: and Biochemistry, Yale University
STREET: 266 Whitney Avenue
CITY: New Haven
STATE: Connecticut
COUNTRY: United States of America
ZIP: 06520-8114
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processing
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,607
FILING DATE: April 27, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB96/01032
FILING DATE: June 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
DESCRIPTION: polypeptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens (melanoma patient immu-
ORGANISM: nized with autologous tumor cells)
INDIVIDUAL ISOLATE: peripheral blood lymphocytes

6140470
SEQ 28

IMMEDIATE SOURCE:
LIBRARY: DM414 scFv antibodies obtained from
LIBRARY: fuses fusion phage construct
CLONE: V13
FEATURE:
NAME/KEY: heavy chain
US-08-983-607-28

Query Match 89.5%; Score 544; DB 3; Length 131;
Best Local Similarity 89.7%; Pred. No. 2e-47;
Matches 104; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 EVOLLEGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSSISGSGSTYY 60
DB 1 EVOLVBSGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSAISGSGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFYDYGQGLTVYSS 116
DB 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKGVAPFDYWGQGLTVYSS 116

RESULT 2

US-09-025-769B-38
Sequence 38, Application US/09025769B
Patent No. 6300064

GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090

INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-769B-38

Query Match 88.7%; Score 539; DB 4; Length 120;
Best Local Similarity 87.5%; Pred. No. 5.7e-47;
Matches 105; Conservative 6; Mismatches 5; Indels 4; Gaps 1;

QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSSISGSGSTYY 60

DB 1 EVOLVBSGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSAISGSGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFYDYGQGLTVYSS 116
DB 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKGVAPFDYWGQGLTVYSS 120

RESULT 3

US-09-025-769B-63
Sequence 63, Application US/09025769B
Patent No. 6300064

GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-769B-63

Query Match 88.7%; Score 539; DB 4; Length 120;
Best Local Similarity 87.5%; Pred. No. 5.7e-47;
Matches 105; Conservative 6; Mismatches 5; Indels 4; Gaps 1;

QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSSISGSGSTYY 60
DB 1 EVOLVBSGGGLVOPGGSLRLSCAAGFTSSFSMSWVRQAPGKLEWVSAISGSGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFYDYGQGLTVYSS 116
DB 61 ADSVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKGVAPFDYWGQGLTVYSS 120

RESULT 4

US-09-025-769B-178
Sequence 178, Application US/09025769B
Patent No. 6300064

GENERAL INFORMATION:
APPLICANT: Knappik, Achim

```

Query Match      88.7%; Score 539; DB 4; Length 281;
Best Local Similarity 87.5%; Pred. No. 1,5e-46;
Matches 105; Conservative 6; Mismatches 5; Indels 4; Gaps 1.

QY      1  EYOLLESGGGLVQPGSLRLSCAASGFTSSFSMSWVRQAPGKLEWVSISGSGSTYY 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      26  EYOLVSSGGGLVQPGGSLRLSCAASGFTSSYMAISWVRQAPGKLEWVSISGSGSTYY 85
QY      61  ADSVKGRFTISRDNSKNTLYIQMNSLRAREDYATYYCAK----PPPFYDWGQSTLYTVSS 116
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      86  ADSVKGRFTISRDNSKNTLYIQMNSLRAREDYATYYCARMGWDGFYADYDWGQSTLYTVSS 145

RESULT 5
US-08-974-899-6
: Sequence 6, Application US/08974899
: Patent No. 6037454
: GENERAL INFORMATION:
: APPLICANT: Presta, Leonard G.
: APPLICANT: Jardieu, Paula M.
: TITLE OF INVENTION: Humanized Anti-CD11a Antibodies
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080

COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

Query Match	88.4%;	Score 537.5;	DB 3;	Length 113;
Best Local Similarity	89.7%;	Pred. No. 7.5e-47;		
Matches 104;	Conservative 5;	Mismatches 4;	Indels 3;	Gaps 1.

QY	1	EVQLLEGGGGLVQPQGSRLRSCASAGFFFSFSMSWRQAPGKLEWVSISGSGSTTY	60
Db	1	EVQLVESGGGLVQPQGSRLRSCASAGFFFSYAMSWRQAPGKLEWVSISDGSGSTY	60
QY	61	ADSVKRRTIFISDNKNTFLYLOMNSLRADENAVYYCAKPPYDPYMGQGLTVVSS	116
Db	61	ADSVKRRPTISDNKNTFLYLOMNSLRADENAVYYCAK---EDYMGQGLTVVSS	113

RESULT 6
 US-08-428-197-1
 Sequence 1, Application US/08428197
 Patent No. 5891438
 GENERAL INFORMATION:
 APPLICANT: SILVERMAN, GREGG J.
 TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
 TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
 TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATIGN AND CONJUGATES
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Spensley Horn Jubas & Lubitz
 STREET: 1880 Century Park East - Suite 500
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90067
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/428,197
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10555
 FILING DATE: 29-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Howells, Stacy L.
 REGISTRATION NUMBER: 34,842
 REFERENCE/DOCKET NUMBER: PD-2630
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 455-5100
 TELEFAX: (619) 455-5110
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 125 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
CLONE: 18/2
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..125
US-08-428-197-1

Query Match 88.4%; Score 537.5; DB 2; Length 125;
Best Local Similarity 84.0%; Pred. No. 8.4e-47;
Matches 105; Conservative 6; Mismatches 5; Indels 9; Gaps 1;

QY 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSVWRQAPGKLEWVSSISGSGSTYY 60
DB 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSVWRQAPGKLEWVSAISGSGSTYY 60
QY 61 ADVKGRFTISRDNKNTLYLQMSLRADTAVYTCAR-----PPPYDYWGQGL 111
DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAVYTCAR-----PPPYDYWGQGL 111
QY 112 VTWSS 116
DB 121 VTWSS 125

RESULT 7

PCT-US93-10555-1
Sequence 1, Application PC/TUS9310555
GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 125 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
CLONE: 18/2
FEATURE:

NAME/KEY: Peptide
LOCATION: 1..125
PCT-US93-10555-1

Query Match 88.4%; Score 537.5; DB 5; Length 125;
Best Local Similarity 84.0%; Pred. No. 8.4e-47;
Matches 105; Conservative 6; Mismatches 5; Indels 9; Gaps 1;

QY 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSVWRQAPGKLEWVSSISGSGSTYY 60
DB 1 EVOLLESGGGLVQPGGSLRLSCAASGFTFSFSMSVWRQAPGKLEWVSAISGSGSTYY 60
QY 61 ADVKGRFTISRDNKNTLYLQMSLRADTAVYTCAR-----PPPYDYWGQGL 111
DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAVYTCAR-----PPPYDYWGQGL 111
QY 112 VTWSS 116
DB 121 VTWSS 125

RESULT 8

US-08-983-607-36
Sequence 36, Application US/08983607
Patent No. 6140470

GENERAL INFORMATION:

APPLICANT: Alan Garen
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
TITLE OF INVENTION: bodies
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Department of Molecular Biophysics
and Biochemistry, Yale University
STREET: 266 Whitney Avenue
CITY: New Haven
STATE: Connecticut
COUNTRY: United States of America
ZIP: 06520-8114

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processing
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,607
FILING DATE: April 27, 1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB96/01032

FILING DATE: June 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maity M. Krinsky
REGISTRATION NUMBER: 32423

REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 116 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
ORIGINAL SOURCE:

ORGANISM: Homo sapiens (melanoma patient immu-
ORGANISM: nized with autologous tumor cells)
INDIVIDUAL ISOLATE: peripheral blood lymphocytes
IMMEDIATE SOURCE:
LIBRARY: DM414 scFv antibodies obtained from

Patent No. 6323322
GENERAL INFORMATION:
APPLICANT: FILPULA, DAVID
APPLICANT: WANG, MAOLIANG
APPLICANT: SHORR, ROBERT
APPLICANT: WHITLOW, MARC
APPLICANT: LEE, LHSYNG S.
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
TITLE OF INVENTION: CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/069,821
FILING DATE: 30-APR-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/067,341
FILING DATE: 02-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-09-069-821-3

Query Match 87.3%; Score 531; DB 4; Length 263;
Best Local Similarity 80.8%; Pred. No. 8.9e-46;
Matches 105; Conservative 7; Mismatches 4; Indels 14; Gaps 2;

QY 1 EVOLLESGGGLVQPGSGLRLSCAASGFTFSFSMSWVRQAPGKGLWVSSISGSS--GTT 58
|||||
DB 130 EVOLVESGGGLVQPGSGLRLSCAASGFTFSFSYAMSVWRQAPGKGLWVSSISGKTIDGST 189
QY 59 YVADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAK-----PPYEDYW 106
|||||
DB 190 YVADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAKRGXSLSGXYYYHYFDYW 249
QY 107 GGGTLVTVSS 116
|||||
DB 250 GGGTLVTVSS 259

RESULT 12
US-09-420-592A-6

Sequence 6, Application US/09420592A
Patent No. 6333396
GENERAL INFORMATION:
APPLICANT: FILPULA, David R.
APPLICANT: Wang, Maoliang
APPLICANT: Whitlow, Marc D.
TITLE OF INVENTION: No. 6333396el Method for Targeted Delivery of Nucleic Acids
FILE REFERENCE: 0977.2300001
CURRENT APPLICATION NUMBER: US/09/420,592A
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/104,949
PRIOR FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 283
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Kabat
NAME/KEY: UNSURE
LOCATION: (232)
OTHER INFORMATION: May be any amino acid.
NAME/KEY: UNSURE
LOCATION: (234)
OTHER INFORMATION: May be any amino acid.
NAME/KEY: UNSURE
LOCATION: (239)
OTHER INFORMATION: May be any amino acid.
US-09-420-592A-6

Query Match 87.3%; Score 531; DB 4; Length 283;
Best Local Similarity 80.8%; Pred. No. 9.7e-46;
Matches 105; Conservative 7; Mismatches 4; Indels 14; Gaps 2;

QY 1 EVOLLESGGGLVQPGSGLRLSCAASGFTFSFSMSWVRQAPGKGLWVSSISGSS--GTT 58
|||||
DB 130 EVOLVESGGGLVQPGSGLRLSCAASGFTFSYAMSVWRQAPGKGLWVSSISGKTIDGST 189
QY 59 YVADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAK-----PPYEDYW 106
|||||
DB 190 YVADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAKRGXSLSGXYYYHYFDYW 249
QY 107 GGGTLVTVSS 116
|||||
DB 250 GGGTLVTVSS 259

RESULT 13
US-10-039-785-53
Sequence 53, Application US/10039785
Patent No. 6538938
GENERAL INFORMATION:
APPLICANT: Salcedo et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
FILE REFERENCE: PF550
CURRENT APPLICATION NUMBER: US/10/039,785
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 60/369,860
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/341,237
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/331,310
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/331,044
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: 60/327,364
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/323,807
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/309,176

PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/294,981
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: 60/293,473
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 53
LENGTH: 249
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: T1006F07 scfv
US-10-039-785-53

Query Match 87.1%; Score 529.5; DB 4; Length 249;
Best Local Similarity 83.3%; Pred. No. 1,2e-45;
Matches 105; Conservative 6; Mismatches 4; Indels 11; Gaps 2;

QY 1 EVQLLESGGGLVQPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWYSSISGSGTYY 60
DB 1 EVQLLESGGGLVQPGGSLRLSCAASGFTSSFSMSWVRQAPGKLEWYSSISGSGSTYY 60
QY 61 ADVSKGRFTISRDNKNTLYIQMNSLRADTAAYVYCAKPPYF-----DYWGQGT 110
DB 61 ADVSKGRFTISRDNKNTLYIQMNSLRADTAAYVYCAK-PSFQMGHYSYGMVWGQGT 119
QY 111 LVTVSS 116
DB 120 MVTVSS 125

RESULT 14
US-07-934-373C-21
Sequence 21, Application US/07934373C
Patent No. 5821337
GENERAL INFORMATION:
APPLICANT: Paul J. Carter
TITLE OF INVENTION: Immunoglobulin Variants
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/934,373C
FILING DATE: 21-Aug-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05126
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids

TYPE: Amino Acid
TOPOLOGY: Linear
US-07-934-373C-21

Query Match 86.8%; Score 528; DB 2; Length 122;
Best Local Similarity 84.4%; Pred. No. 7.3e-46;
Matches 103; Conservative 6; Mismatches 7; Indels 6; Gaps 1;

QY 1 EVQLLESGGGLVQPGGSLRLSCAASGFTSSFSMSWVRQAPGKLEWYSSISGSGTYY 60
DB 1 EVQLVLESGGGLVQPGGSLRLSCAASGFTSSFSMSWVRQAPGKLEWYSSISGSGSTYY 60
QY 61 ADVSKGRFTISRDNKNTLYIQMNSLRADTAAYVYCAKPPYF-----FPIFYDYGQGLTVTV 114
DB 61 ADVSKGRFTISRDNKNTLYIQMNSLRADTAAYVYCARGVYSLGLVDYWGQGLTVTV 120
QY 115 SS 116
DB 121 SS 122

RESULT 15
US-08-437-642B-21
Sequence 21, Application US/08437642B
Patent No. 6054297
GENERAL INFORMATION:
APPLICANT: Paul J. Carter
TITLE OF INVENTION: Immunoglobulin Variants
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,642B
FILING DATE: 09-May-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/934373
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/146206
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05126
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-437-642B-21

Query Match 86.8%; Score 528; DB 3; Length 122;

Best Local Similarity 84.4%; Pred. No. 7.3e-46;
Matches 103; Conservative 6; Mismatches 7; Indels - 6; Gaps 1;

```
QY      1 EVQLLESGGGLVQPQGGSLRLSCAASGFTFSFSMSGWRQAPGKGLEWVSSISGSSGTTY 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 EVQLVESGGGLVQPQGGSLRLSCAASGFTFSYAMSQWRQAPGKGLEWVSVISGDGTTY 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      61 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCAKP-----FPYEDYWGQGLVTV 114
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCARGVGYSLGLIDYWGQGLVTV 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      115 SS 116
      |||
Db      121 SS 122
```

Search completed: August 20, 2003, 12:44:16
Job time : 81.3614 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 4.04819 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-20

Perfect score: 73

Sequence: 1 GDGSSGGSGGASTG 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/pdata/1/paa/US06_NEW_COMB.pep:*
2: /cgn2_6/pdata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/pdata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/pdata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/pdata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/pdata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/pdata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	238	US-10-336-041A-1	Sequence 1, Appl
2	73	100.0	307	US-10-291-265-332	Sequence 332, App
3	73	100.0	312	US-10-291-265-334	Sequence 334, App
4	73	100.0	384	US-10-291-265-804	Sequence 804, App
5	73	100.0	384	US-10-291-265-805	Sequence 805, App
6	73	100.0	384	US-10-291-265-806	Sequence 806, App
7	73	100.0	384	US-10-291-265-807	Sequence 807, App
8	62	84.9	240	US-10-336-041A-10	Sequence 10, Appl
9	62	84.9	240	US-10-336-041A-11	Sequence 11, Appl
10	62	84.9	241	US-10-336-041A-13	Sequence 13, Appl
11	62	84.9	247	US-10-336-041A-9	Sequence 9, Appl
12	62	84.9	211	US-10-273-973-16	Sequence 16, Appl
13	52	71.2	433	US-60-485-404-55	Sequence 55, Appl
14	52	71.2	498	US-10-225-838B-21	Sequence 21, Appl
15	52	71.2	1079	US-09-820-843B-20	Sequence 20, Appl
16	52	71.2	16	US-10-350-555-72	Sequence 72, Appl
17	51	69.9	276	US-10-292-798-1694	Sequence 1694, Ap
18	51	69.9	329	US-10-282-798-1438	Sequence 1438, Ap
19	50	68.5	329	US-10-282-798-1438	Sequence 1438, Ap
20	50	68.5	340	US-60-480-890-2711	Sequence 2711, Ap
21	50	68.5	2161	US-10-408-765A-1283	Sequence 1283, Ap
22	48	65.8	202	US-60-490-890-2252	Sequence 2252, Ap
23	48	65.8	249	US-10-603-113-27128	Sequence 27128, A
24	48	65.8	256	US-09-820-843B-18	Sequence 18, Appl
25	48	65.8	347	US-60-487-610-1610	Sequence 1610, Ap
26	48	65.8	446	US-10-447-243-3	Sequence 3, Appl

27	48	65.8	484	US-09-820-843B-19	Sequence 19, Appl
28	48	65.8	561	US-60-490-890-1788	Sequence 1788, Ap
29	48	65.8	634	US-10-374-780A-1466	Sequence 1466, Ap
30	48	65.8	821	US-10-292-798-1370	Sequence 1370, Ap
31	48	65.8	1487	US-10-273-573-9934	Sequence 9934, Ap
32	48	65.8	1487	US-10-408-765A-744	Sequence 744, Appl
33	47	64.4	34	PCT-US02-31375-82	Sequence 82, Appl
34	47	64.4	93	US-10-612-783-6563	Sequence 6563, Ap
35	47	64.4	570	US-10-603-113-20916	Sequence 20916, A
36	47	64.4	620	US-10-286-897-6628	Sequence 6628, Ap
37	47	64.4	620	US-10-258-898A-3056	Sequence 3056, Ap
38	47	64.4	627	US-10-286-897-3056	Sequence 3056, Ap
39	47	64.4	627	US-10-258-898A-3056	Sequence 1527, Ap
40	47	64.4	637	US-60-487-610-1527	Sequence 795, App
41	46.5	63.0	254	US-10-374-780A-795	Sequence 5411, Ap
42	46	63.0	87	US-10-612-783-5411	Sequence 5714, Ap
43	46	63.0	134	US-10-286-897-5714	Sequence 5714, Ap
44	46	63.0	134	US-10-258-898A-5714	Sequence 5714, Ap
45	46	63.0	292	US-10-374-780A-2080	Sequence 2080, Ap

ALIGNMENTS

```

RESULT 1
US-10-336-041A-1
; Sequence 1, Application US/10336041A
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P_WOAS US/10/336, 041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (117)..(130)
; OTHER INFORMATION: Linker
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (131)..(238)
; OTHER INFORMATION: VL
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
; OTHER INFORMATION: antibody fragment
US-10-336-041A-1

Query Match      100.0%; Score 73; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GDGSSGGSGGASTG 14
Db      117 GDGSSGGSGGASTG 130

RESULT 2
US-10-291-265-332
; Sequence 332, Application US/10291265
; GENERAL INFORMATION:

```

```
;; APPLICANT: Hyseq, Inc.
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT APPLICATION NUMBER: US/10/291,265
;; CURRENT FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/631,451
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 332
;; LENGTH: 307
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-332
```

```
Query Match      100.0%; Score 73; DB 6; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GDGSSGSGSGASTG 14
        |||
Db       136 GDGSSGSGSGASTG 149
```

```
RESULT 3
US-10-291-265-334
;; Sequence 334, Application US/10291265
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT APPLICATION NUMBER: US/10/291,265
;; CURRENT FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/631,451
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 334
;; LENGTH: 312
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-334
```

```
Query Match      100.0%; Score 73; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GDGSSGSGSGASTG 14
        |||
Db       141 GDGSSGSGSGASTG 154
```

```
RESULT 4
US-10-291-265-804
;; Sequence 804, Application US/10291265
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
```

```
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT APPLICATION NUMBER: US/10/291,265
;; CURRENT FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/631,451
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 804
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-804
```

```
Query Match      100.0%; Score 73; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GDGSSGSGSGASTG 14
        |||
Db       156 GDGSSGSGSGASTG 169
```

```
RESULT 5
US-10-291-265-805
;; Sequence 805, Application US/10291265
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT APPLICATION NUMBER: US/10/291,265
;; CURRENT FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/631,451
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 805
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-805
```

```
Query Match      100.0%; Score 73; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GDGSSGSGSGASTG 14
        |||
Db       156 GDGSSGSGSGASTG 169
```

```
RESULT 6
US-10-291-265-806
;; Sequence 806, Application US/10291265
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT APPLICATION NUMBER: US/10/291,265
;; CURRENT FILING DATE: 2000-01-25
```

```
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/631,451
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 806
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-806

Query Match          100.0%; Score 73; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GDGSSGSGSGASTG 14
Db      156 GDGSSGSGSGASTG 169

RESULT 7
US-10-291-265-807
;; Sequence 807, Application US/10291265
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-017 (785)
;; CURRENT FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: US/10/291,265
;; PRIOR FILING DATE: 2000-01-25
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 09/617,746
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 09/633,870
;; PRIOR FILING DATE: 2000-09-15
;; NUMBER OF SEQ ID NOS: 944
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 807
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-291-265-807

Query Match          100.0%; Score 73; DB 6; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GDGSSGSGSGASTG 14
Db      156 GDGSSGSGSGASTG 169

RESULT 8
US-10-336-041A-10
;; Sequence 10, Application US/10336041A
;; GENERAL INFORMATION:
;; APPLICANT: Schering AG
;; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
;; FILE REFERENCE: 27041P.WOAS
;; CURRENT FILING DATE: 2003-01-03
;; PRIOR APPLICATION NUMBER: US/10/336,041A
;; PRIOR FILING DATE: 2003-01-03
;; PRIOR APPLICATION NUMBER: EP02 000 315.8
;; PRIOR FILING DATE: 2002-02-25
;; PRIOR APPLICATION NUMBER: US60/358702
;; PRIOR FILING DATE: 2002-02-25
```

```
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10
;; LENGTH: 240
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant
US-10-336-041A-10

Query Match          84.9%; Score 62; DB 6; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GDGSSGSGSGAS 12
Db      117 GDGSSGSGSGAS 128

RESULT 9
US-10-336-041A-12
;; Sequence 12, Application US/10336041A
;; GENERAL INFORMATION:
;; APPLICANT: Schering AG
;; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
;; FILE REFERENCE: 27041P.WOAS
;; CURRENT FILING DATE: 2003-01-03
;; PRIOR APPLICATION NUMBER: US/10/336,041A
;; PRIOR FILING DATE: 2003-01-03
;; PRIOR APPLICATION NUMBER: EP02 000 315.8
;; PRIOR FILING DATE: 2002-01-03
;; PRIOR APPLICATION NUMBER: US60/358702
;; PRIOR FILING DATE: 2002-02-25
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 12
;; LENGTH: 240
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant
US-10-336-041A-12

Query Match          84.9%; Score 62; DB 6; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GDGSSGSGSGAS 12
Db      117 GDGSSGSGSGAS 128

RESULT 10
US-10-336-041A-11
;; Sequence 11, Application US/10336041A
;; GENERAL INFORMATION:
;; APPLICANT: Schering AG
;; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
;; FILE REFERENCE: 27041P.WOAS
;; CURRENT FILING DATE: 2003-01-03
;; PRIOR APPLICATION NUMBER: US/10/336,041A
;; PRIOR FILING DATE: 2002-01-03
;; PRIOR APPLICATION NUMBER: EP02 000 315.8
;; PRIOR FILING DATE: 2002-02-25
;; PRIOR APPLICATION NUMBER: US60/358702
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 241
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
```

OTHER INFORMATION: Description of Artificial Sequence: recombinant
OTHER INFORMATION: antibody fragment
US-10-336-041A-11

Query Match 84.9%; Score 62; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGSGAS 12
DB 117 GDGSSGSGSGAS 128

RESULT 11
US-10-336-041A-13
Sequence 13, Application US/10336041A
GENERAL INFORMATION:
APPLICANT: Schering AG
TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
FILE REFERENCE: 27041P-WOAS
CURRENT APPLICATION NUMBER: US/10/336, 041A
CURRENT FILING DATE: 2003-01-03
PRIOR APPLICATION NUMBER: EP02 000 315.8
PRIOR FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: US60/358702
PRIOR FILING DATE: 2002-02-25
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 241
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant
US-10-336-041A-13

Query Match 84.9%; Score 62; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGSGAS 12
DB 117 GDGSSGSGSGAS 128

RESULT 12
US-10-336-041A-9
Sequence 9, Application US/10336041A
GENERAL INFORMATION:
APPLICANT: Schering AG
TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
FILE REFERENCE: 27041P-WOAS
CURRENT APPLICATION NUMBER: US/10/336, 041A
CURRENT FILING DATE: 2003-01-03
PRIOR APPLICATION NUMBER: EP02 000 315.8
PRIOR FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: US60/358702
PRIOR FILING DATE: 2002-02-25
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 247
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant
US-10-336-041A-9

Query Match 84.9%; Score 62; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGSGAS 12
DB 117 GDGSSGSGSGAS 128

RESULT 13
US-10-273-973-16
Sequence 16, Application US/10273973
GENERAL INFORMATION:
APPLICANT: Kang, Angray
Barbas, Carlos
Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of Patent Counsel
STREET: 10666 North Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/273, 973
FILING DATE: 22-Jan-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/133, 011
FILING DATE: 08-JUN-1994
APPLICATION NUMBER: US 07/683, 602
FILING DATE: 10-APR-1991
APPLICATION NUMBER: US 07/826, 623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34, 163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-273-973-16

Query Match 71.2%; Score 52; DB 6; Length 211;
Best Local Similarity 71.4%; Pred. No. 7.7;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDGSSGSGSGAS 14
DB 23 GDGSSGSGSGSGS 36

RESULT 14
US-60-485-404-55
Sequence 55, Application US/60485404
GENERAL INFORMATION:

```

; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homayoun
; TITLE OF INVENTION: Transferrin Fusion Protein Libraries
; FILE REFERENCE: 054710-5007-PR
; CURRENT APPLICATION NUMBER: US/60/485,404
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US 10/231,494
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/406,997
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 10/384,060
; PRIOR FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 60/460,829
; PRIOR FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 433
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: M13 pIII coat protein for insertion into pUC18 vectors
US-60-485-404-55

```

```

Query Match          71.2%: Score 52; DB 7; Length 433;
Best Local Similarity 71.4%: Pred. No. 15;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY      1 GDGSSGGSGGSGASTG 14
        ||| ||||| ||
Db      245 GGGSGGSGGSGGSEG 258

```

```

RESULT 15
US-10-225-838B-21
; Sequence 21, Application US/10225838B
; GENERAL INFORMATION:
; APPLICANT: Neuronz, Ltd.
; APPLICANT: Neuronz Biosciences, Inc.
; APPLICANT: Sleg, Frank
; APPLICANT: Hughes, Paul
; TITLE OF INVENTION: Neural Regeneration Peptides and Methods for Their Use In
; FILE REFERENCE: NRNZ-1023US1
; CURRENT APPLICATION NUMBER: US/10/225,838B
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: 60/314,952
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-225-838B-21

```

```

Query Match          71.2%: Score 52; DB 6; Length 498;
Best Local Similarity 64.3%: Pred. No. 17;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY      1 GDGSSGGSGGSGASTG 14
        ||| ||||| ||
Db      434 GDGSGGAGAGNGTG 447

```

Search completed: August 20, 2003, 12:45:11
 Job time : 4.04819 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 ; Search time 30.7831 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-20

Perfect score: 73

Sequence: 1 GDSSSGSGGASTG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.19Jun03.*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	14	21	AAV53773
2	73	100.0	14	22	AAE08818
3	73	100.0	238	21	AAV53775
4	73	100.0	307	22	AAU14225
5	73	100.0	312	22	AAU14227
6	73	100.0	384	22	AAW24101
7	73	100.0	384	22	AAU14461
8	73	100.0	384	22	AAU14462
9	73	100.0	384	22	AAU14463

10	73	100.0	384	22	AAU14464	Human novel protei
11	55	75.3	928	21	ABB92408	Herbically activ
12	54	74.0	318	23	ABB81209	Mycobacterium tube
13	53	72.6	296	20	AAV14121	Raly protein seque
14	52	71.2	35	18	AAW17096	Phage-derived flex
15	52	71.2	36	16	AAW79295	Glyser-rich spacer
16	52	71.2	36	16	AAW93972	Highly flexible (G
17	52	71.2	36	20	AAV02130	Peptide linker use
18	52	71.2	36	22	AAE13901	Chemically modifie
19	52	71.2	36	23	ABG97743	Human interleukin-
20	52	71.2	36	23	ABG65749	Linker sequence us
21	52	71.2	49	22	AAE14027	Chemically modifie
22	52	71.2	49	22	ABG97800	Human interleukin-
23	52	71.2	60	16	AAW65750	Linker sequence us
24	52	71.2	60	22	AAE14028	Chemically modifie
25	52	71.2	60	23	ABG97801	Human interleukin-
26	52	71.2	211	15	AAW62927	Bacteriophage coat
27	52	71.2	211	15	AAW50174	Membrane anchor pe
28	52	71.2	211	15	AAW54281	M13 phage coat pro
29	52	71.2	211	21	AAW95082	Amino acid seq ID
30	52	71.2	211	21	AAW98191	Amino acid seq ID
31	52	71.2	219	22	AAV97677	Protein from const
32	52	71.2	219	22	AAV97686	hEGF/M13 pIII fusi
33	52	71.2	233	16	AAW75660	hEGF-M13 pIII fusi
34	52	71.2	233	22	AAV97794	M13pAg-EGF-pIII fu
35	52	71.2	233	22	AAV97790	IL-3 containing fu
36	52	71.2	301	16	AAW79318	Human interleukin-
37	52	71.2	301	21	AAV53199	Human interleukin-
38	52	71.2	301	22	AAE13993	Chemically modifie
39	52	71.2	301	23	ABG97766	Human interleukin-
40	52	71.2	319	23	ABW81746	Region #1 of bacte
41	52	71.2	334	21	ABW81230	Mycobacterium boyl
42	52	71.2	335	16	AAW79319	IL-3 containing fu
43	52	71.2	335	21	AAV53200	Human interleukin-
44	52	71.2	335	22	AAE13994	Myeloprotein (MPO)
45	52	71.2	335	23	ABG97767	Human interleukin-

ALIGNMENTS

RESULT 1
AAV53773 standard; Peptide: 14 AA.

AAV53773; 22-FEB-2000 (first entry)

Linker of an antibody with improved specificity for fibronectin.

scFv; antibody; ED-B domain epitope; fibronectin; marker;
angiogenesis; vascular proliferation; diabetic retinopathy;
age-related macular degeneration; tumour; immunosuppressive detection;
blood coagulation; blood vessel occlusion; ocular angiogenesis;
angiogenesis-related pathology.

Synthetic.

WO9958570-A2.

18-NOV-1999.

11-MAY-1999; 99WO-EP03210.

11-MAY-1998; 98US-0075338.

28-APR-1999; 99US-0300425.

(EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.

Neri D, Tarli L, Vitti F, Birchler M;

WPI; 2000-039074/03.

XX Fibronectin ED-B domain epitope specific antibodies and conjugate
 PT antibodies
 XX
 PS Claim 10; Page 38; 59pp; English.
 XX
 CC The present sequence represents a linker component of a modified human
 CC scFv antibody which has specific affinity for a characteristic epitope
 CC of the ED-B domain of fibronectin. The affinity of the antibody for
 CC this epitope was improved by introducing a number of mutations in the
 CC complementarity determining region (CDR) residues located at the
 CC periphery of the binding site. The improved antibody is used for rapid
 CC targeting markers of angiogenesis, for detecting diseases characterized
 CC by vascular proliferation or tumours. The antibody localizes the respective
 CC tissue within 3 to 4 hours after injection. It is used in
 CC immunocytographic detection of angiogenesis and for diagnosis and
 CC therapy of tumours and diseases characterized by vascular proliferation.
 CC The antibody can be conjugated to a molecule which induces blood
 CC coagulation and blood vessel occlusion. These conjugates are used in
 CC the preparation of injectable compositions for the treatment of
 CC angiogenesis-related pathologies, especially caused by or associated
 CC with ocular angiogenesis.
 CC
 SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 21; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDSSSGSGSGASTG 14
 |||||
 Db 1 GDSSSGSGSGASTG 14

RESULT 2
 AA08818
 ID AA08818 standard; Protein: 14 AA.
 XX
 AC AA08818;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Synthetic peptide linker for constructing scFv L19 antibody.
 XX
 KW scFv; single-chain variable antibody fragment; cancer; cytotoxic;
 KW coagulant; ED-B domain; fibronectin; tumour; ocular disorder; psoriasis;
 KW vascular proliferation; rheumatoid arthritis; blood vessel occlusion;
 KW angiogenesis; blood coagulation.
 XX
 OS Synthetic.
 XX
 PN W0200162800-A1.
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001MO-EP02062.
 XX
 PR 24-FEB-2000; 2000US-0512082.
 XX
 PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
 XX
 PI Neri D, Tarli L, Viti F, Birchler M;
 XX
 PS WPI; 2001-541701/60.
 DR
 XX
 PT An antibody, with specific affinity for a characteristic epitope of the
 PT ED-B domain of fibronectin for the treatment of diseases characterized
 PT by vascular proliferation
 XX
 PS Claim 10; Page 36; 73pp; English.
 XX
 CC The invention relates to an antibody with specific affinity for a

CC characteristic epitope of the ED-B domain of fibronectin, where the
 CC antibody has improved affinity to ED-B. The invention also relates to
 CC conjugates comprising antibodies with a suitable photoactive molecule
 CC useful in the detection and/or coagulation of blood vessels. An antibody
 CC with improved affinity to the ED-B domain is useful for diagnosis and
 CC therapy of tumours and diseases characterised by vascular proliferation,
 CC cancer, rheumatoid arthritis, neo-vasculature associated ocular
 CC disorders and psoriasis. Treatment of angiogenesis related pathologies
 CC comprises the injection of conjugates comprising antibody and a molecule
 CC capable of inducing blood coagulation and blood vessel occlusion. The
 CC present sequence is a peptide linker used for constructing single-chain
 CC variable antibody fragment (scFv) L19 antibody related to the invention.
 CC
 SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDSSSGSGSGASTG 14
 |||||
 Db 1 GDSSSGSGSGASTG 14

RESULT 3
 AA53775
 ID AA53775 standard; Protein: 238 AA.
 XX
 AC AA53775;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE An antibody with improved specificity for fibronectin.
 XX
 KW scFv; antibody; ED-B domain epitope; fibronectin; marker;
 KW angiogenesis; vascular proliferation; diabetic retinopathy;
 KW age-related macular degeneration; tumour; immunocytographic detection;
 KW blood coagulation; blood vessel occlusion; ocular angiogenesis;
 KW angiogenesis-related pathology.
 XX
 OS Synthetic.
 XX
 OS Homo sapiens.
 XX
 PN W0958570-A2.
 PD 18-NOV-1999.
 XX
 PF 11-MAY-1999; 99WO-EP03210.
 XX
 PR 11-MAY-1998; 98US-0075338.
 PR 28-APR-1999; 99US-0300425.
 XX
 PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
 XX
 PI Neri D, Tarli L, Viti F, Birchler M;
 XX
 PS WPI; 2000-039074/03.
 DR
 XX
 PT Fibronectin ED-B domain epitope specific antibodies and conjugate
 PT antibodies
 XX
 PS Claim 10; Page 7; 59pp; English.
 XX
 CC The present sequence represents a modified human scFv antibody which
 CC has specific affinity for a characteristic epitope of the ED-B domain
 CC of fibronectin. The affinity of the antibody for this epitope was
 CC improved by introducing a number of mutations in the complementarity
 CC determining region (CDR) residues located at the periphery of the
 CC binding site. The improved antibody is used for rapid targeting markers
 CC of angiogenesis, for detecting diseases characterized by vascular
 CC proliferation, such as diabetic retinopathy, age-related macular
 CC degeneration or tumours. The antibody localizes the respective tissue
 CC within 3 to 4 hours after injection. It is used in immunocytographic

CC detection of angiogenesis and for diagnosis and therapy of tumours and
 CC diseases characterized by vascular proliferation. The antibody can be
 CC conjugated to a molecule which induces blood coagulation and blood
 CC vessel occlusion. These conjugates are used in the preparation of
 CC injectable compositions for the treatment of angiogenesis-related
 CC pathologies, especially caused by or associated with ocular
 CC angiogenesis.
 CC note: this sequence does not appear in the specification: it is an
 CC amalgamation of the sequences given in claim 10.
 CC
 SO Sequence 238 AA:
 Query Match 100.0%; Score 73; DB 21; Length 238;
 Best Local Similarity 100.0%; Pred. No. 0.084;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GDGSSGGSGGASTG 14
 117 GDGSSGGSGGASTG 130
 Db 117 GDGSSGGSGGASTG 130

RESULT 4
 AAU14225
 ID AAU14225 standard; Protein: 307 AA.
 AC AAU14225;
 XX 24-OCT-2001 (first entry)
 DT 24-OCT-2001 (first entry)
 DE Human novel protein #96.
 XX
 DE Human; novel protein; Antianemic; osteopathic; antiinflammatory;
 KW Immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200155437-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02623.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-451939/48.
 XX
 DR N-PSDB: AAS22530.
 XX
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 XX Example 4; Page 575; 894pp; English.
 XX
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to

CC raise antibodies/elicite an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.
 CC
 SO Sequence 307 AA:
 Query Match 100.0%; Score 73; DB 22; Length 307;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GDGSSGGSGGASTG 14
 136 GDGSSGGSGGASTG 149
 Db 136 GDGSSGGSGGASTG 149

RESULT 5
 AAU14227
 ID AAU14227 standard; Protein: 312 AA.
 AC AAU14227;
 XX 24-OCT-2001 (first entry)
 DT 24-OCT-2001 (first entry)
 DE Human novel protein #98.
 XX
 DE Human; novel protein; Antianemic; osteopathic; antiinflammatory;
 KW Immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200155437-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02623.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-451939/48.
 XX
 DR N-PSDB: AAS22532.
 XX
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 XX Example 4; Page 576-577; 894pp; English.
 XX
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and

CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicits an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

CC Sequence 312 AA;

Query Match 100.0%; Score 73; DB 22; Length 312;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDSSGGSGGASTG 14
 |||||
 Db 141 GDSSGGSGGASTG 154

RESULT 6
 AAM24101
 ID AAM24101 standard; Protein: 384 AA.

XX AAM24101;

XX 12-OCT-2001 (first entry)

XX Human EST encoded protein SEQ ID NO: 1626.

XX Human: sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.

XX Homo sapiens.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001MO-US02687.

XX 25-JAN-2000; 2000US-0491404.

XX 17-JUL-2000; 2000US-0617746.

XX 03-AUG-2000; 2000US-0631451.

XX 15-SEP-2000; 2000US-0663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

XX Cao Y, Drmanac RA, Zhang J, Werhman T;

XX WPI; 2001-476164/51.

XX N-PSDB; AAB98760.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising

XX antibodies and research use -

XX Claim 20; Page 1102-1103; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel

XX proteins from a variety of organisms, including human, dog, cat, horse,

XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea

CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.

XX Sequence 384 AA;

Query Match 100.0%; Score 73; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDSSGGSGGASTG 14
 |||||
 Db 153 GDSSGGSGGASTG 166

RESULT 7
 AAU14461
 ID AAU14461 standard; Protein: 384 AA.

XX AAU14461;

XX 24-OCT-2001 (first entry)

XX Human novel protein #332.

XX Human: novel protein; Antianemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.

XX Homo sapiens.

XX WO200155437-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001MO-US02623.

XX 25-JAN-2000; 2000US-0491404.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-451939/48.

XX N-PSDB; AAS22766.

XX Isolated polypeptides useful for treating anti-inflammatory diseases,

XX nervous system disorders, and for regenerating bone and cartilage -

XX Example 4; Page 824-825; 894pp; English.

XX The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicits an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,

CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

SO Sequence 384 AA;

Query Match 100.0%; Score 73; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. NO. 0.14;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDGSSGGSGASTG 14
 |||||
 Db 156 GDGSSGGSGASTG 169

RESULT 8
 AAU14462

ID AAU14462 standard; Protein: 384 AA.

AC AAU14462;

DT 24-OCT-2001 (first entry)

DE Human novel protein #333.

Human: novel protein; Antianemic; osteopathic; antiinflammatory;
 immunomodulatory; cytosolic; neuroprotective; vulnery; nootropic;
 anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
 thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 tissue regeneration; immune disorder.

OS Homo sapiens.

PN WO200155437-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02623.

PR 25-JAN-2000; 2000US-0491404.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-451939/48.

DR N-PSDB; AAS22767.

PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 PS Example 4; Page 825-826; 894pp; English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to

CC raise antibodies/elicite an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

SO Sequence 384 AA;

Query Match 100.0%; Score 73; DB 22; Length 384;
 Best Local Similarity 100.0%; Pred. NO. 0.14;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDGSSGGSGASTG 14
 |||||
 Db 156 GDGSSGGSGASTG 169

RESULT 9
 AAU14463

ID AAU14463 standard; Protein: 384 AA.

AC AAU14463;

DT 24-OCT-2001 (first entry)

DE Human novel protein #334.

Human: novel protein; Antianemic; osteopathic; antiinflammatory;
 immunomodulatory; cytosolic; neuroprotective; vulnery; nootropic;
 anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
 thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 tissue regeneration; immune disorder.

OS Homo sapiens.

PN WO200155437-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02623.

PR 25-JAN-2000; 2000US-0491404.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-451939/48.

DR N-PSDB; AAS22768.

PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 PS Example 4; Page 826-827; 894pp; English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and

CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

XX Sequence 384 AA:

Query Match 100.0%; Score 73; DB 22; Length 384;

Best Local Similarity 100.0%; Pred. No. 0.14; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDSSSGSGGASTG 14
 DB 156 GDSSSGSGGASTG 169

RESULT 10
 AAU14464
 ID AAU14464 standard; Protein; 384 AA.

AC AAU14464;
 XX 24-OCT-2001 (first entry)

DE Human novel protein #335.

KW Human; novel protein; Antianemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytostatic; neuroprotective; vulnerray; nocotropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antitumoral; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antidiabetic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.

OS Homo sapiens.

PN WO20015437-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02623.

PR 25-JAN-2000; 2000US-0491404.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-451939/48.

DR N-PSDB; AAS22769.

PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -

PS Example 4; Page 827; 894pp; English.

XX The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant

CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

XX Sequence 384 AA:

Query Match 100.0%; Score 73; DB 22; Length 384;

Best Local Similarity 100.0%; Pred. No. 0.14; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDSSSGSGGASTG 14
 DB 156 GDSSSGSGGASTG 169

RESULT 11
 ABB92408
 ID ABB92408 standard; Protein; 928 AA.

AC ABB92408;

DE 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1619.

KW Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

PS

XX Claim 5; SEQ ID NO 1619; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value

CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

SO Sequence 928 AA;

Query Match 75.3%; Score 55; DB 23; Length 928;
Best Local Similarity 78.6%; Pred. No. 59;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GDGSSGSGGASTG 14
|||||:|||||
DB 448 GDGSSPTGTGASG 461

RESULT 12

ABB81229
ID ABB81229 standard; Protein: 318 AA.

XX ABB81229;

XX 16-AUG-2002 (first entry)

DE Mycobacterium tuberculosis strain H37Rv PE-PGRS protein sequence.

KW Mycobacterium tuberculosis; Mycobacterium bovis; mycobacterium;

KM detection; BAC vector; bacterial artificial chromosome; tuberculosis.

OS Mycobacterium tuberculosis.

XX WO954487-A2.

PN 28-OCT-1999.

XX 16-APR-1999; 99WO-IB00740.

XX 16-APR-1998; 98US-0060756.

PA (INSP) INST PASTEUR.

PI Cole S, Buchrieser-Brosch R, Gordon S, Billault A;

DR WPI: 2000-013262/01.

PT Isolation of polynucleotides from mycobacterial genomes, useful for
XX detection of Mycobacteria and for combating tuberculosis -

PS Disclosure; Fig 6; 161pp; English.

CC The present invention describes a method for isolating a polynucleotide
CC of interest that is present or is expressed in a genome of a first
CC mycobacterium strain and that is absent or altered in a genome of a
CC second mycobacterium strain, which is different from the first strain.
CC using a bacterial artificial chromosome (BAC) vector. Recombinant BAC
CC vectors, which are preferably immobillised, can be used to detect
CC mycobacterial nucleic acids (genomic DNA, cDNA or mRNA) in biological
CC samples. The polynucleotides identified are useful as probes or primers
CC for detecting a given mycobacterium of interest. By aligning the
CC polynucleotides contained in the recombinant BAC vectors it is possible
CC to physically map a polynucleotide of mycobacterial origin in a
CC biological sample. The methods and vectors from the present invention
CC are useful in providing information for combating tuberculosis. It is
CC possible to compare genomes between different strains or species and
CC their non-pathogenic strains or species counterparts. AB062492 to
CC AB063228 and ABB81227 to ABB81230 represent sequences used in the
CC exemplification of the present invention.

SO Sequence 318 AA;

Query Match 74.0%; Score 54; DB 21; Length 318;
Best Local Similarity 76.9%; Pred. No. 27;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GDGSSGSGGAST 13
|||:|||||
DB 88 GDGAGAGIGGAST 100

RESULT 13

AAV14121
ID AAV14121 standard; Protein: 296 AA.

XX AAV14121;

XX 21-JUL-1999 (first entry)

DE Raly protein sequence.

KW PCR primer; specific engineered fragment; gene targeting vector;
KM yeast selectable marker; mammalian cell selectable marker;
XX gene function identification.

OS Synthetic.

PN WO9923239-A1.

XX 14-MAY-1999.

XX 03-JUN-1998; 98WO-US11388.

XX 31-OCT-1997; 97US-0963602.

PA (AMGE-) AMGEN INC.

PI Garfinkel D, Woychik R;

DR WPI: 1999-313353/26.

DR N-PSDB; AAX58236.

PT New gene targeting vectors

PS Example 6; Fig 8; 66pp; English.

CC This sequence is the Raly (RNP associated with lethal yellow) protein,
CC DNA encoding it can be used in a specific engineered fragment (SEF).
CC The invention relates to gene targeting vectors that comprise 2 yeast
CC selectable markers; a mammalian cell selectable marker and gene-specific
CC sequences for homologous recombination. The gene targeting vectors are
CC comprising a first yeast selectable marker, a bacterial shuttle vector
CC marker, and a fragment of genomic DNA containing at least part of a gene
CC to be targeted; (b) preparing a SEF comprising a marker cassette, the
CC marker cassette comprising a second yeast selectable marker different
CC from the first yeast selectable marker, a selectable marker capable of
CC expression in mammalian embryonic stem cells and the marker cassette
CC being flanked on each side by mammalian gene-specific flanking sequences
CC homologous to a portion of the gene to be targeted; (c) transforming
CC yeast cells with a shuttle vector as in (a) and with a SEF as in (b), and
CC allowing the shuttle vector and the SEF to recombine by homologous
CC recombination; (d) selecting the transformed yeast cells for expression
CC of the first and second yeast selectable markers; and (e) isolating the
CC targeting vector produced by recombination between the shuttle vector and
CC the SEF from the yeast cells selected in (d). The method can be used for
CC knocking out or inserting genes or regulatory sequences into preselected
CC genetic loci. The method can be used for identification of the function
CC of genes and their roles in diseases. Using the method, detailed mapping
CC and sequence information are not required in order to prepare targeting
CC constructs or restriction sites which results in a significant saving of
CC time and effort in preparing targeting constructs.

SO Sequence 296 AA;

Query Match 72.6%; Score 53; DB 20; Length 296;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 727
LENGTH: 318
TYPE: PRT
ORGANISM: Mycobacterium sp.
US-09-670-314-727

Query Match
Best Local Similarity 74.0%; Score 54; DB 4; Length 318;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGSGGASTG 13
DB 88 GGGSGGSGGASTG 100

RESULT 3
US-08-470-775-11
Sequence 11, Application US/08470775
Patent No. 5858347

GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Maitre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Paik, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Co-administration of Interleukin-3
TITLE OF INVENTION: Mutants with Colony Stimulating Factors
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.,
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,775
FILING DATE: 06-JUN-1995

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2788/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-470-775-11

Query Match
Best Local Similarity 71.2%; Score 52; DB 2; Length 36;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGSGGSGGASTG 14
DB 1 GGGSGGSGGSGSEG 14

RESULT 4
US-08-469-318-50
Sequence 50, Application US/08469318
Patent No. 6022535

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,318
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,872
FILING DATE:
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-469-318-50

Query Match
Best Local Similarity 71.2%; Score 52; DB 3; Length 36;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGSGGSGGASTG 14
DB 1 GGGSGGSGGSGSEG 14

RESULT 5
US-08-468-609A-50
Sequence 50, Application US/08468609A
Patent No. 6030812

GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Maitre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Paik, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (I
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.,
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,609A
FILING DATE: 06-JUN-1995

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192,325
FILING DATE: 14-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2790/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6986
TELEFAX: (314)737-6972
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-468-609A-50

Query Match 71.2%; Score 52; DB 3; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GDSSGGSGGASTG 14
| | | | | | | | | | | | | | | | | |
Db 1 GGSGGGSGGSGGSEG 14

RESULT 6
US-08-192-325B-50

Sequence 50, Application US/08192325B
Patent No. 6057133
GENERAL INFORMATION:

APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Bradford-Goldberg, Sarah R.
APPLICANT: Caparon, Maïre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Palk, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Multivalent IL-3 Hematopoiesis
TITLE OF INVENTION: Fusion Protein
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/192,325B
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2790
TELECOMMUNICATION INFORMATION:

TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881

INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:

LENGTH: 36 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-192-325B-50

Query Match 71.2%; Score 52; DB 3; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GDSSGGSGGASTG 14
| | | | | | | | | | | | | | | | | |
Db 1 GGSGGGSGGSGGSEG 14

RESULT 7
US-08-559-267A-11

Sequence 11, Application US/08559267A
Patent No. 6074639
GENERAL INFORMATION:

APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Bradford-Goldberg, Sarah R.
APPLICANT: Caparon, Maïre H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Palk, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Ex-vivo expansion of hematopoietic cells using Interleukin-
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESSEE: Corporate Patent Dept.
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/559,267A
FILING DATE: 15-NOV-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192,299
FILING DATE: 04-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2788/4
TELECOMMUNICATION INFORMATION:

TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-559-267A-11

Query Match 71.2%; Score 52; DB 3; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GDSSGGSGGASTG 14
| | | | | | | | | | | | | | | | | |

Db 1 GGGSGGGSGGSGSEG 14

RESULT 8
US-08-469-124-11
Sequence 11, Application US/08469124
Patent No. 6132991
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Mair H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Palk, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Co-administration of Interleukin-3
TITLE OF INVENTION: Mutants with Colony Stimulating Factors
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,124
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2788/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6881
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-469-124-11

Query Match 71.2%; Score 52; DB 3; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGSGTG 14
Db 1 GGGSGGGSGGSGSEG 14

RESULT 9
US-08-446-872A-50
Sequence 50, Application US/08446872A
Patent No. 6361977
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Mair H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.

APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Palk, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
TITLE OF INVENTION: Fusion Protein
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,872A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192,325
FILING DATE: 14-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2790/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6986
TELEFAX: (314)737-6972
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-872A-50

Query Match 71.2%; Score 52; DB 4; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGSGTG 14
Db 1 GGGSGGGSGGSGSEG 14

RESULT 10
US-08-762-227A-50
Sequence 50, Application US/08762227A
Patent No. 6436387
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Mair H.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Palk, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
TITLE OF INVENTION: Fusion Protein
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
STREET: P. O. Box 5110

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,227A
FILING DATE: 09-DEC-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192,325
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: US 08/446,872
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2790/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-08-762-227A-50

Query Match 71.2%; Score 52; DB 4; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9; 4; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GDGSSGGSGGSG 14
DB 1 GGGSGGGSGGSG 14

RESULT 11
PCT-US95-01185-50
Sequence 50, Application PC/TUS9501185
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01185
FILING DATE: 02-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192325
FILING DATE: 14-FEB-1994
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-01185-50

Query Match 71.2%; Score 52; DB 5; Length 36;
Best Local Similarity 71.4%; Pred. No. 1.9;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GDGSSGGSGGSG 14
DB 1 GGGSGGGSGGSG 14

RESULT 12
US-08-469-318-194
Sequence 194, Application US/08469318
Patent No. 6022535
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,318
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,872
FILING DATE:
INFORMATION FOR SEQ ID NO: 194:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-469-318-194



Query Match 71.2%; Score 52; DB 3; Length 49;
Best Local Similarity 71.4%; Pred. No. 2.5;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGSG 14
DB 10 GGGSGGGSGGSG 23

RESULT 13
US-08-468-609A-194
Sequence 194, Application US/08468609A
Patent No. 6030812
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Bratford-Goldberg, Sarah R.
APPLICANT: Caparon, Alan M.
APPLICANT: Easton, Alan M.
APPLICANT: Klein, Barbara K.
APPLICANT: Mckean, John P.
APPLICANT: Olin, Peter O.
APPLICANT: Paik, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (I
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
ADDRESS: Corporate Patent Dept.,
STREET: P. O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

1  COMPUTER:  IBM PC compatible
2  OPERATING SYSTEM:  PC-DOS/MS-DOS
3  SOFTWARE:  PatentIn Release #1.0, Version #1.25
4  CURRENT APPLICATION DATA:
5  APPLICATION NUMBER:  US/08/468, 609A
6  FILING DATE:  06-JUN-1995
7  CLASSIFICATION:
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER:  US 08/192,325
10 FILING DATE:  14-FEB-1994
11 ATTORNEY/AGENT INFORMATION:
12 NAME:  Bennett, Dennis A.
13 REGISTRATION NUMBER:  34,547
14 REFERENCE/DOCKET NUMBER:  C-2790/3
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE:  (314)737-6986
17 TELEFAX:  (314)737-6972
18 INFORMATION FOR SEQ ID NO:  194:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH:  49 amino acids
21 TYPE:  amino acid
22 STRANDEDNESS:
23 TOPOLOGY:  linear
24 MOLECULE TYPE:  peptide
25 US-08-468-609A-194

```

SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 194:
US-08-762-227A-194

Query Match 71.2% Score 52; DB 4; Length 49;
Best Local Similarity 71.4% Pred. No. 2.5;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GDGSSGSGSGASTG 14
1 1111111111
Db 10 GCGSGGSGSGSGSEG 23

Search completed: August 20, 2003, 12:44:17
Job time : 10.6988 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 ; Search time 13.1928 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-27

Sequence: 1 A1SGSG 6

AISGSG 6

Scoring table: BLOSUM62

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum	Match	100%
74.4	54.4	45

Listing first 45 summaries

Database :

1. A_Geneseq_13Jun03: *

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11980.DAT: *
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11981.DAT: *
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11982.DAT: *
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11983.DAT: *
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11984.DAT: *
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11985.DAT: *
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11986.DAT: *
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11987.DAT: *
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11988.DAT: *
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11989.DAT: *
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11990.DAT: *
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11991.DAT: *
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11992.DAT: *
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11993.DAT: *
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11994.DAT: *
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11995.DAT: *
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11996.DAT: *
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11997.DAT: *
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11998.DAT: *
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A11999.DAT: *
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A12000.DAT: *
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A12001.DAT: *
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A12002.DAT: *
24:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/A12003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	100.0	10	20	AAV39626	CTLA-4 VLD CDR 100c
2	28	100.0	17	18	AAV16651	Anti-cancer specific
3	28	100.0	17	21	AAV95213	Anti-platelet glyco
4	28	100.0	17	22	AAV65311	Anti-IL-18 antibod
5	28	100.0	17	22	AAV65317	Anti-IL-18 antibod
6	28	100.0	17	24	ABR01548	Human anti-TIMP-1
7	28	100.0	25	20	AAV39564	CTLA-4 VLD CDR re
8	28	100.0	56	20	AAV02512	clone selected after
9	28	100.0	56	20	AAV02516	clone selected after

[illegible]

ALIGNMENTS

RESULT 1	
AAV39626	
ID	AAV39626 standard; peptide; 10 AA

KM CTNA-4; cytotoxic T-lymphocyte associated antigen 4; binding agent; VLD
variable-like domain; human; diagnosis; cancer; blood clot.

US
OS

PN W09945110-A1

PD 10-SEP-1999.

PF 05-MAR-1999; 99WO-AU00136.

PR 06-MAR-1998; 98AU-0002210.

PA (DIAT-) DIATECH PTY LTD.

PI Cola G, Galanis M, Hudson PJ, Irving RA, Nuttall SD;

DR. WPI; 1999-551040/46.

PT New binding agent comprising monomeric V-like domain in which at least one complementarity determining region loop is modified, useful for

PT diagnosis of cancer -
 XX
 PS Example 6; Fig 6; 117pp; English.
 CC This sequence is variable-like domain (VLD) of the human cytotoxic
 CC T-lymphocyte associated antigen 4 (CTLA-4) CDR loop replacement, used
 CC in the binding agent of the invention. The binding agent (1) comprises at
 CC least one monomeric VLD that is derived from a non-antibody ligand and
 CC has at least one CDR (complementarity determining region) loop sequence,
 CC or part of it, modified or replaced so that, compared to unmodified VLD,
 CC its solubility is increased and/or the size is altered and/or a
 CC disulphide bond is created within, or between, one or more CDR loops. (1)
 CC are used for diagnosis, e.g. in vivo detection/localisation of cancer,
 CC blood clots etc., also in vitro when immobilised on solid supports or
 CC biosensors and therapeutically. Modified VLD may have binding affinity
 CC for drugs, steroids, pesticides, antigens, growth factors, tumour
 CC markers, cell or viral proteins. Modification of VLD improves solubility
 CC and alters binding specificity. Since VLD are derived from human
 CC proteins, the need for a humanizing step (to avoid adverse immune
 CC responses) is avoided, and modification also improves expression in
 CC Escherichia coli.
 CC
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 28; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A1SGSG 6
 Db 1 A1SGSG 6
 RESULT 2
 ID AAM16651 standard; Peptide; 17 AA.
 AC AAM16651;
 DT 29-JAN-1998 (first entry)
 DE Anti-cancer specific antigen Mab heavy chain CDR2.
 DE Heavy chain; variable region; cancer specific antigen; human;
 KW monoclonal antibody; hMab; diagnosis; cancer; immunotherapy;
 KM purification; complementarity determining region 2; CDR2.
 OS Homo sapiens.
 XX JP09098786-A.
 PN 15-APR-1997.
 PD 06-OCT-1995; 95JP-0284400.
 PF 06-OCT-1995; 95JP-0284400.
 PR 06-OCT-1995; 95JP-0284400.
 XX (MOMI) MORINAGA & CO LTD.
 PA (SHKJ) SHINGIJUTSU JIGYODAN.
 XX WPI: 1997-375445/25.
 DR N-PSDB; AAT66783.
 XX
 PT cDNA encoding human monoclonal antibody - useful in medicine, or to
 PT purify cancer specific antigen
 PS Claim 9; Fig 12; 7pp; Japanese.
 CC The present sequence is the heavy chain complementarity
 CC determining region 2 of an anti-cancer specific antigen human
 CC monoclonal antibody (hMab). The hMab can be used in medicine, e.g.
 CC clinical diagnosis of cancer or immunotherapy, or to purify cancer
 CC specific antigen. The industrial scale production of large amounts

CC of the hMab is made feasible by genetic engineering using the hMab
 CC cDNA.
 SQ Sequence 17 AA;
 Query Match 100.0%; Score 28; DB 18; Length 17;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A1SGSG 6
 Db 1 A1SGSG 6
 RESULT 3
 ID AAY95213 standard; Peptide; 17 AA.
 AC AAY95213;
 DT 29-AUG-2000 (first entry)
 DE Anti-platelet glycoprotein Ib human H1b-3 VH CDR2.
 DE Variable heavy chain; single chain antibody; scFv; human; H1b-3;
 KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
 KM antithrombotic; thrombus; therapy; diagnostic; CDR2;
 KW complementarity determining region.
 OS Homo sapiens.
 XX WO200026667-A1.
 PN 11-MAY-2000.
 PD 29-OCT-1999; 99WO-US25495.
 PF 30-OCT-1998; 98US-0106275.
 PR (MILL) MILLER J L.
 PA MILLER JL;
 PI
 PT WPI: 2000-365744/31.
 DR Isolated nucleic acid molecule encoding anti-human platelet
 XX glycoprotein Ib alpha molecule useful for producing antibodies which
 PT inhibit platelet aggregation -
 PS Claim 15; Fig 7; 89pp; English.
 CC The present sequence is that of complementarity determining region
 CC 2 (CDR2) of the heavy chain variable region (VH) of human
 CC single chain antibody (scFv) H1b-3 (see AAY95213), which is directed
 CC against platelet glycoprotein Ib (GPIb). The H1b series of scFv
 CC was isolated from a human synthetic VH and VL scFv library on the
 CC basis of their binding to platelet GPIb. Whether displayed as
 CC surface proteins on a phagemid or secreted as free scFv by
 CC Escherichia coli, the H1b scFv clones are capable of inhibiting
 CC von Willebrand factor-dependent aggregation of platelets. The scFv
 CC are composed of native human protein sequences and are therefore
 CC attractive potential reagents for therapeutic purposes. They
 CC provide a new class of antithrombotic agents, useful for the
 CC prevention of platelet-dependent thrombi in diseased arteries,
 CC bypass grafts, dialysis etc., and can also be used as diagnostic
 CC reagents. Methods of inhibiting aggregation of platelets, of
 CC binding human platelet GPIb alpha and of selecting a VH or VL
 CC region of an antibody that inhibits platelet aggregation are
 CC claimed. Fragments of the scFv VH or VL chain, including CDR
 CC fragments, are also claimed.
 SQ Sequence 17 AA;

Query Match 100.0%; Score 28; DB 21; Length 17;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 1 AISGSG 6

RESULT 4
 AAG65311

ID AAG65311 standard; protein: 17 AA.

AC AAG65311;

DT 30-NOV-2001 (first entry)

DE Anti-IL-18 antibody LT28 heavy chain CDR2 fragment.

KW IL-18; interleukin-18; human; antibody; antirheumatic; cerebroprotective; neurotropic; neurological; antiinflammatory; antiparkinsonian; cardiant; immunosuppressive; antidepressant; neuroleptic; hepatotropic; LT28.

XX Homo sapiens.

OS WO200158956-A2.

PN 16-AUG-2001.

PP 09-FEB-2001; 2001WO-US04170.

PR 10-FEB-2000; 2000US-0181608.

PI (BADI) BASF AG.

PA Ghayur T, Dixon RW, Roguska M, White M, Labkovsky B, Salfeld J;

PI Duncan AR, Brocklehurst SM, Mankovich J, Shorrock CP, Thompson JE;

PI Lennard SN;

DR WPI: 2001-550020/61.

XX Novel antibodies and compounds capable of binding to human
 XX interleukin-18 useful for treating, e.g., inflammatory disorders,
 XX neurological disorders, heart failure, myocardial infarction, and
 XX autoimmune diseases -

PS Claim 32; Page 41; 91pp; English.

XX The invention provides isolated antibodies, or antigen-binding portions,
 XX that are capable of binding to human interleukin-18 (IL-18). The
 XX antibodies may be used to inhibit human IL-18 activity in, and treat a
 XX disorder where IL-18 is detrimental in, a human subject suffering from,
 XX inflammatory disorders (e.g., rheumatoid arthritis, Crohn's disease,
 XX inflammatory bowel disease, and osteoarthritis), neurological disorders
 XX (e.g., Huntington's chorea, Parkinson's disease, Alzheimer's disease, and
 XX stroke), heart failure, myocardial infarction, autoimmune diseases such
 XX as autoimmune hepatitis and autoimmune neutropenia, and mental disorders
 XX (e.g., depression and schizophrenia). Treatment with an anti-IL-18
 XX antibody may occur before, concurrent, or after administration of a
 XX second agent selected from an antibody, or fragment, capable of binding
 XX human IL-12, methotrexate, anti-tumor necrosis factor, corticosteroids,
 XX cyclosporin, rapamycin, FK506, and non-steroidal anti-inflammatory
 XX agents. The present sequence represents an anti-IL-18 antibody LT28
 XX heavy chain CDR2 fragment.

SO Sequence 17 AA;

Query Match 100.0%; Score 28; DB 22; Length 17;

Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||

DB 1 AISGSG 6

RESULT 5
 AAG65317

ID AAG65317 standard; protein: 17 AA.

AC AAG65317;

DT 30-NOV-2001 (first entry)

DE Anti-IL-18 antibody heavy chain variable region fragment.

KW IL-18; interleukin-18; human; antibody; antirheumatic; cerebroprotective; neurotropic; neurological; antiinflammatory; antiparkinsonian; cardiant; immunosuppressive; antidepressant; neuroleptic; hepatotropic; LT28.

XX Homo sapiens.

OS WO200158956-A2.

PN 16-AUG-2001.

PP 09-FEB-2001; 2001WO-US04170.

PR 10-FEB-2000; 2000US-0181608.

PI (BADI) BASF AG.

PA Ghayur T, Dixon RW, Roguska M, White M, Labkovsky B, Salfeld J;

PI Duncan AR, Brocklehurst SM, Mankovich J, Shorrock CP, Thompson JE;

PI Lennard SN;

DR WPI: 2001-550020/61.

XX Novel antibodies and compounds capable of binding to human
 XX interleukin-18 useful for treating, e.g., inflammatory disorders,
 XX neurological disorders, heart failure, myocardial infarction, and
 XX autoimmune diseases -

PS Claim 36; Page 78; 91pp; English.

XX The invention provides isolated antibodies, or antigen-binding portions,
 XX that are capable of binding to human interleukin-18 (IL-18). The
 XX antibodies may be used to inhibit human IL-18 activity in, and treat a
 XX disorder where IL-18 is detrimental in, a human subject suffering from,
 XX inflammatory disorders (e.g., rheumatoid arthritis, Crohn's disease,
 XX inflammatory bowel disease, and osteoarthritis), neurological disorders
 XX (e.g., Huntington's chorea, Parkinson's disease, Alzheimer's disease, and
 XX stroke), heart failure, myocardial infarction, autoimmune diseases such
 XX as autoimmune hepatitis and autoimmune neutropenia, and mental disorders
 XX (e.g., depression and schizophrenia). Treatment with an anti-IL-18
 XX antibody may occur before, concurrent, or after administration of a
 XX second agent selected from an antibody, or fragment, capable of binding
 XX human IL-12, methotrexate, anti-tumor necrosis factor, corticosteroids,
 XX cyclosporin, rapamycin, FK506, and non-steroidal anti-inflammatory
 XX agents. The present sequence represents an anti-IL-18 antibody heavy
 XX chain variable region fragment.

SO Sequence 17 AA;

Query Match 100.0%; Score 28; DB 22; Length 17;

Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 1 AISGSG 6

RESULT 6

ABR01548
 ID ABR01548 standard; peptide: 17 AA.

```

XX ABR01548;
AC 16-APR-2003 (first entry)
DE
DT
XX
XX Human anti-TIMP-1 Fab VHCDR2 #1.
DE
XX
XX Human; antibody; tissue inhibitor of metalloproteinase-1; TIMP-1; VHCDR3;
XX matrix metalloproteinase; MMP; variable heavy chain; VHCDR3; hepatotropic;
XX variable light chain; cytosolic; nephrotropic; cardiac; liver fibrosis;
XX alcoholic liver disease; cardiac fibrosis; acute coronary syndrome;
XX lupus nephritis; glomerulosclerotic renal disease; lung cancer; VHCDR1;
XX idiopathic pulmonary fibrosis; benign prostate hypertrophy; colon cancer.
XX
XX Homo sapiens.
XX
XX WO200286085-A2.
XX
XX 31-OCT-2002.
XX
XX 24-APR-2002; 2002WO-US12801.
XX
XX 24-APR-2001; 2001US-285683P.
XX
XX (FARB) BAYER CORP.
XX (MORP-) MORPHOSYS AG.
XX
XX Pan C, Knorr AM, Schauer M, Hirth-dietrich C, Kraft S, Krebs B;
XX WPI; 2003-129114/12.
XX
XX New human anti-TIMP-1 (tissue inhibitor of metalloproteinase-1)
XX antibodies, for diagnosing or ameliorating the symptoms of a disorder
XX in which TIMP-1 is elevated, e.g. liver fibrosis, benign prostate
XX hypertrophy or lung cancer
XX
XX Example 17; Page 47; 228pp; English.
XX
XX The invention relates to a novel purified preparation of a human
XX antibody, which binds to a tissue inhibitor of metalloproteinase-1 (TIMP-1)
XX and neutralises a matrix metalloproteinase (MMP)-inhibiting activity of
XX TIMP-1. The antibody comprises a variable heavy chain (VHC)DR3 region and
XX a variable light chain (VLC)DR3 region. An antibody preparation of the
XX invention has hepatotropic, cytosolic, nephrotropic and cardiac
XX activity. The human antibody is useful for decreasing an MMP-inhibiting
XX activity of a TIMP-1. It is especially useful for ameliorating the
XX symptoms of a disorder in which TIMP-1 is elevated, e.g. liver
XX fibrosis, alcoholic liver disease, cardiac fibrosis, acute coronary
XX syndrome, lupus nephritis, glomerulosclerotic renal disease, idiopathic
XX pulmonary fibrosis, benign prostate hypertrophy, lung cancer or colon
XX cancer. The antibody is also useful for detecting a TIMP-1 in a test
XX preparation, or in diagnosing a disorder in which a TIMP-1 level is
XX elevated. The present sequence represents a VHCDR1 region of an antibody.
XX
XX Sequence 17 AA;
XX
XX Query Match 100.0%; Score 28; DB 24; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 61;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AISGSG 6
XX |
XX |
XX |
XX |
XX |
XX |
XX 1 AISGSG 6
XX
XX Db
XX
XX RESULT 7
XX AAY39564
XX ID AAY39564 standard; peptide; 25 AA.
XX
XX AC AAY39564;
XX
XX XX 23-NOV-1999 (first entry)

```

```

XX CTLA-4 VLD CDR2 region variant.
DE
XX
XX CTLA-4; cytotoxic T-lymphocyte associated antigen 4; binding agent; VLD;
XX variable-like domain; human; diagnosis; cancer; blood clot.
XX
XX Homo sapiens.
XX
XX WO9945110-A1.
XX
XX 10-SEP-1999.
XX
XX 05-MAR-1999; 99WO-AU00136.
XX
XX 06-MAR-1998; 98AU-0002210.
XX
XX (DIAT-) DIATECH PTY LTD.
XX
XX Cola G, Galanis M, Hudson PJ, Irving RA, Nuttall SD;
XX WPI; 1999-551040/46.
XX
XX New binding agent comprising monomeric V-like domain in which at least
XX one complementarity determining region loop is modified, useful for
XX diagnosis of cancer
XX
XX Example 1; Page 18; 117pp; English.
XX
XX This sequence is a variant of the CDR2 of the variable-like domain (VLD)
XX of the human cytotoxic T-lymphocyte associated antigen 4 (CTLA-4), used
XX in the binding agent of the invention. The binding agent (I) comprises at
XX least one monomeric VLD that is derived from a non-antibody ligand and
XX has at least one CDR (complementarity determining region) loop sequence,
XX or part of it, modified or replaced so that, compared to unmodified VLD,
XX its solubility is increased and/or the size is altered and/or a
XX disulphide bond is created within, or between, one or more CDR loops. (I)
XX are used for diagnosis, e.g. in vivo detection/localisation of cancer,
XX blood clots etc., also in vitro when immobilised on solid supports or
XX biosensors and therapeutically. Modified VLD may have binding affinity
XX for drugs, steroids, pesticides, antigens, growth factors, tumour
XX markers, cell or viral proteins. Modification of VLD improves solubility
XX and alters binding specificity. Since VLD are derived from human
XX proteins, the need for a humanizing step (to avoid adverse immune
XX responses) is avoided, and modification also improves expression in
XX Escherichia coli.
XX
XX Sequence 25 AA;
XX
XX Query Match 100.0%; Score 28; DB 20; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 89;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AISGSG 6
XX |
XX |
XX |
XX |
XX |
XX |
XX 9 AISGSG 14
XX
XX Db
XX
XX RESULT 8
XX AAY02512
XX ID AAY02512 standard; Protein; 56 AA.
XX
XX AC AAY02512;
XX
XX XX 15-JUL-1999 (first entry)
XX
XX DE Clone selected after panning a NNK library of the invention.
XX
XX Screening; functional polypeptide; ligand; non-functional;
XX enrichment; single chain antibody; Scfv.
XX
XX OS Synthetic.
XX
XX XX WO9920749-A1.

```

```

XX 29-APR-1999.
PD
XX 20-OCT-1998; 98WO-GB03135.
XX
XX 21-NOV-1997; 97US-0066729.
XX 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX Tomlinson I, Winter G;
XX WPI, 1999-288302/24.
XX
XX Screening for functional polypeptides which bind a ligand
XX
XX Example 3; Fig 4; 67pp; English.
XX
XX The specification describes a method for screening for functional
XX polypeptides which bind a ligand. The method comprises contacting a
XX repertoire of polypeptides with a generic ligand, and then screening
XX selected functional polypeptides with a target ligand. The method
XX permits the removal from a chosen repertoire of polypeptides, those
XX which are non-functional, e.g. as a result of the introduction of
XX frame-shift mutations, stop codons, folding mutants or expression
XX mutants which would be or are incapable of binding to any target
XX ligand. The method also permits the enrichment of a chosen repertoire
XX of polypeptides for those polypeptides which are functional, well folded
XX and highly expressed. The polypeptides obtained can be used in
XX diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
XX represent clones selected after panning primary and somatic NKK
XX libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
XX NIP-BSA and hen egg lysozyme).
XX
XX Sequence 56 AA:
XX
XX Query Match 100.0%; Score 28; DB 20; Length 56;
XX Best Local Similarity 100.0%; Pred. No. 2e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 AISGSG 6
DB 6 AISGSG 11

```

```

PI Tomlinson I, Winter G;
XX
XX WPI, 1999-288302/24.
XX
XX Screening for functional polypeptides which bind a ligand
XX
XX Example 3; Fig 4; 67pp; English.
XX
XX The specification describes a method for screening for functional
XX polypeptides which bind a ligand. The method comprises contacting a
XX repertoire of polypeptides with a generic ligand, and then screening
XX selected functional polypeptides with a target ligand. The method
XX permits the removal from a chosen repertoire of polypeptides, those
XX which are non-functional, e.g. as a result of the introduction of
XX frame-shift mutations, stop codons, folding mutants or expression
XX mutants which would be or are incapable of binding to any target
XX ligand. The method also permits the enrichment of a chosen repertoire
XX of polypeptides for those polypeptides which are functional, well folded
XX and highly expressed. The polypeptides obtained can be used in
XX diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
XX represent clones selected after panning primary and somatic NKK
XX libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
XX NIP-BSA and hen egg lysozyme).
XX
XX Sequence 56 AA:
XX
XX Query Match 100.0%; Score 28; DB 20; Length 56;
XX Best Local Similarity 100.0%; Pred. No. 2e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 AISGSG 6
DB 6 AISGSG 11

```

```

RESULT 10
AAY02517
ID AAY02517 standard; Protein: 56 AA.
XX
XX AAY02517:
XX
XX 15-JUL-1999 (first entry)
XX
XX Clone selected after panning a NKK library of the invention.
XX
XX Screening: functional polypeptide; ligand: non-functional;
XX enrichment: single chain antibody; ScFv.
XX
XX Synthetic.
XX
XX WO9920749-A1.
XX
XX 29-APR-1999.
XX
XX 20-OCT-1998; 98WO-GB03135.
XX
XX 21-NOV-1997; 97US-0066729.
XX 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX
XX Tomlinson I, Winter G;
XX
XX WPI, 1999-288302/24.
XX
XX Screening for functional polypeptides which bind a ligand
XX
XX Example 3; Fig 4; 67pp; English.
XX
XX The specification describes a method for screening for functional
XX polypeptides which bind a ligand. The method comprises contacting a
XX repertoire of polypeptides with a generic ligand, and then screening

```

CC selected functional polypeptides with a target ligand. The method
CC permits the removal from a chosen repertoire of polypeptides, those
CC which are non-functional, e.g. as a result of the introduction of
CC frame-shift mutations, stop codons, folding mutants or expression
CC mutants which would be or are incapable of binding to any target
CC ligand. The method also permits the enrichment of a chosen repertoire
CC of polypeptides for those polypeptides which are functional, well folded
CC and highly expressed. The polypeptides obtained can be used in
CC diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
CC represent clones selected after panning primary and somatic NK
CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
CC NIP-BSA and hen egg lysozyme).

XX Sequence 56 AA;

Query Match 100.0%; Score 28; DB 20; Length 56;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
|||||
DB 6 AISGSG 11

RESULT 11
AAY02518
ID AAY02518 standard; Protein; 56 AA.

XX AAY02518;

DT 15-JUL-1999 (first entry)

XX Clone selected after panning a NK library of the invention.

KW Screening; functional polypeptide; ligand: non-functional;
KM enrichment; single chain antibody; scfv.

XX Synthetic.

XX WO9920749-A1.

XX 29-APR-1999.

XX 20-OCT-1998; 98WO-GB03135.

XX 21-NOV-1997; 97US-0066729.

XX 20-OCT-1997; 97GB-0022131.

XX 13-NOV-1997; 97US-0065428.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Tomlinson I, Winter G;

XX WPI, 1999-288302/24.

XX Screening for functional polypeptides which bind a ligand

XX Example 3; Fig 4; 67pp; English.

XX The specification describes a method for screening for functional

XX polypeptides which bind a ligand. The method comprises contacting a

XX repertoire of polypeptides with a generic ligand, and then screening

XX selected functional polypeptides with a target ligand. The method

XX permits the removal from a chosen repertoire of polypeptides, those

XX which are non-functional, e.g. as a result of the introduction of

XX frame-shift mutations, stop codons, folding mutants or expression

XX mutants which would be or are incapable of binding to any target

XX ligand. The method also permits the enrichment of a chosen repertoire

XX of polypeptides for those polypeptides which are functional, well folded

XX and highly expressed. The polypeptides obtained can be used in

XX diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525

XX represent clones selected after panning primary and somatic NK

XX libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,

XX NIP-BSA and hen egg lysozyme).

CC NIP-BSA and hen egg lysozyme).

XX Sequence 56 AA;

Query Match 100.0%; Score 28; DB 20; Length 56;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
|||||
DB 6 AISGSG 11

RESULT 12
AAY02520
ID AAY02520 standard; Protein; 56 AA.

XX AAY02520;

DT 15-JUL-1999 (first entry)

XX Clone selected after panning a NK library of the invention.

KW Screening; functional polypeptide; ligand: non-functional;
KM enrichment; single chain antibody; scfv.

XX Synthetic.

XX WO9920749-A1.

XX 29-APR-1999.

XX 20-OCT-1998; 98WO-GB03135.

XX 21-NOV-1997; 97US-0066729.

XX 20-OCT-1997; 97GB-0022131.

XX 13-NOV-1997; 97US-0065428.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Tomlinson I, Winter G;

XX WPI, 1999-288302/24.

XX Screening for functional polypeptides which bind a ligand

XX Example 3; Fig 4; 67pp; English.

XX The specification describes a method for screening for functional

XX polypeptides which bind a ligand. The method comprises contacting a

XX repertoire of polypeptides with a generic ligand, and then screening

XX selected functional polypeptides with a target ligand. The method

XX permits the removal from a chosen repertoire of polypeptides, those

XX which are non-functional, e.g. as a result of the introduction of

XX frame-shift mutations, stop codons, folding mutants or expression

XX mutants which would be or are incapable of binding to any target

XX ligand. The method also permits the enrichment of a chosen repertoire

XX of polypeptides for those polypeptides which are functional, well folded

XX and highly expressed. The polypeptides obtained can be used in

XX diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525

XX represent clones selected after panning primary and somatic NK

XX libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,

XX NIP-BSA and hen egg lysozyme).

RESULT 13
AAV02522
ID AAV02522 standard; Protein: 56 AA.
XX
AC AAV02522;
XX
DT 15-JUL-1999 (first entry)
XX
DE Clone selected after panning a NNK library of the invention.
XX
KW Screening; functional polypeptide; ligand; non-functional;
KW enrichment; single chain antibody; scfv.
XX
OS Synthetic.
XX
PN WO9920749-A1.
XX
PD 29-APR-1999.
XX
PE 20-OCT-1998; 98WO-GB03135.
XX
PR 21-NOV-1997; 97US-0066729.
PR 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Tomlinson I, Winter G;
XX
DR WPI; 1999-288302/24.
XX
PT Screening for functional polypeptides which bind a ligand
XX
PS Example 3; Fig 4; 67pp; English.
XX
CC The specification describes a method for screening for functional
CC polypeptides which bind a ligand. The method comprises contacting a
CC repertoire of polypeptides with a generic ligand, and then screening
CC selected functional polypeptides with a target ligand. The method
CC permits the removal from a chosen repertoire of polypeptides, those
CC which are non-functional, e.g. as a result of the introduction of
CC frame-shift mutations, stop codons, folding mutants or expression
CC mutants which would be or are incapable of binding to any target
CC ligand. The method also permits the enrichment of a chosen repertoire
CC of polypeptides for those polypeptides which are functional, well folded
CC and highly expressed. The polypeptides obtained can be used in
CC diagnostic, prophylactic and therapeutic procedures. AAV02473-102525
CC represent clones selected after panning primary and somatic NNK
CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
CC NIP-BSA and hen egg lysozyme).
XX
SO Sequence 56 AA:
Query Match 100.0%; Score 28; DB 20; Length 56;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
DB 6 AISGSG 11
IIIIII
IIIIII

RESULT 14
AAV02524
ID AAV02524 standard; Protein: 56 AA.
XX
AC AAV02524;
XX
DT 15-JUL-1999 (first entry)
XX
DE Clone selected after panning a NNK library of the invention.

XX
KW Screening; functional polypeptide; ligand; non-functional;
KW enrichment; single chain antibody; scfv.
XX
OS Synthetic.
XX
PN WO9920749-A1.
XX
PD 29-APR-1999.
XX
PE 20-OCT-1998; 98WO-GB03135.
XX
PR 21-NOV-1997; 97US-0066729.
PR 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Tomlinson I, Winter G;
XX
DR WPI; 1999-288302/24.
XX
PT Screening for functional polypeptides which bind a ligand
XX
PS Example 3; Fig 4; 67pp; English.
XX
CC The specification describes a method for screening for functional
CC polypeptides which bind a ligand. The method comprises contacting a
CC repertoire of polypeptides with a generic ligand, and then screening
CC selected functional polypeptides with a target ligand. The method
CC permits the removal from a chosen repertoire of polypeptides, those
CC which are non-functional, e.g. as a result of the introduction of
CC frame-shift mutations, stop codons, folding mutants or expression
CC mutants which would be or are incapable of binding to any target
CC ligand. The method also permits the enrichment of a chosen repertoire
CC of polypeptides for those polypeptides which are functional, well folded
CC and highly expressed. The polypeptides obtained can be used in
CC diagnostic, prophylactic and therapeutic procedures. AAV02473-102525
CC represent clones selected after panning primary and somatic NNK
CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
CC NIP-BSA and hen egg lysozyme).
XX
SO Sequence 56 AA:
Query Match 100.0%; Score 28; DB 20; Length 56;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
DB 6 AISGSG 11
IIIIII
IIIIII

RESULT 15
AAV02525
ID AAV02525 standard; Protein: 56 AA.
XX
AC AAV02525;
XX
DT 15-JUL-1999 (first entry)
XX
DE Clone selected after panning a NNK library of the invention.
XX
KW Screening; functional polypeptide; ligand; non-functional;
KW enrichment; single chain antibody; scfv.
XX
OS Synthetic.
XX
PN WO9920749-A1.
XX
PD 29-APR-1999.
XX
PE 20-OCT-1998; 98WO-GB03135.

XX 21-NOV-1997; 97US-0066729.
 PR 20-OCT-1997; 97GB-0022131.
 PR 13-NOV-1997; 97US-0065428.
 XX

PA (MEDT-) MEDICAL RES COUNCIL.

XX Tomlinson I, Winter G;
 PI

XX WPI; 1999-288302/24.
 DR

XX Screening for functional polypeptides which bind a ligand
 PT

XX Example 3; Fig 4; 67pp; English.
 PS

XX The specification describes a method for screening for functional
 CC polypeptides which bind a ligand. The method comprises contacting a
 CC repertoire of polypeptides with a generic ligand, and then screening
 CC selected functional polypeptides with a target ligand. The method
 CC permits the removal from a chosen repertoire of polypeptides, those
 CC which are non-functional, e.g. as a result of the introduction of
 CC frame-shift mutations, stop codons, folding mutants or expression
 CC mutants which would be or are incapable of binding to any target
 CC ligand. The method also permits the enrichment of a chosen repertoire
 CC of polypeptides for those polypeptides which are functional, well folded
 CC and highly expressed. The polypeptides obtained can be used in
 CC diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
 CC represent clones selected after panning primary and somatic NMK
 CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
 CC NIP-BSA and hen egg lysozyme).
 XX

SQ Sequence 56 AA;

Query Match 100.0%; Score 28; DB 20; Length 56;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 Db 6 AISGSG 11

Search completed: August 20, 2003, 12:33:41
 Job time : 14.1928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41; Search time 4.15663 Seconds
(without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-27
Perfect score: 28
Sequence: 1 A1SGSG 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTC05.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	98	2	US-08-428-197-48
2	28	100.0	98	5	PCT-US93-10555-48
3	28	100.0	116	2	US-08-428-197-2
4	28	100.0	116	3	US-08-983-607-36
5	28	100.0	116	5	PCT-US93-10555-2
6	28	100.0	117	3	US-08-545-809A-109
7	28	100.0	117	3	US-08-983-607-46
8	28	100.0	118	2	US-08-652-816A-12
9	28	100.0	119	3	US-08-963-607-22
10	28	100.0	120	4	US-09-025-769B-38
11	28	100.0	122	2	US-08-958-201-6
12	28	100.0	122	2	US-08-428-197-38
13	28	100.0	123	2	US-08-428-197-38
14	28	100.0	123	5	PCT-US93-10555-38
15	28	100.0	124	3	US-08-983-607-51
16	28	100.0	125	1	US-08-478-039-99
17	28	100.0	125	1	US-08-476-349A-99
18	28	100.0	125	5	PCT-US93-10555-1
19	28	100.0	125	5	US-08-983-607-28
20	28	100.0	131	3	US-08-983-607-32
21	28	100.0	140	3	US-08-983-607-32
22	28	100.0	144	1	US-08-026-320A-2
23	28	100.0	249	4	US-09-025-769B-178
24	28	100.0	281	4	US-09-025-769B-178
25	28	100.0	450	4	US-09-171-337A-5
26	28	100.0	461	4	US-09-631-022-5
27	28	100.0	461	4	US-09-631-022-5

28	100.0	1218	4	US-09-589-567-2	Sequence 2, Appl
29	96.4	3281	4	US-09-328-352-8170	Sequence 8170, Ap
30	96.4	329	4	US-08-887-534A-24	Sequence 24, Appl
31	96.4	329	4	US-09-527-431-24	Sequence 24, Appl
32	96.4	332	4	US-09-134-001C-5229	Sequence 5229, Ap
33	96.4	338	4	US-09-107-532A-6222	Sequence 6222, Ap
34	96.4	351	4	US-08-311-731A-28	Sequence 28, Appl
35	96.4	452	4	US-09-252-991A-20578	Sequence 20578, A
36	96.4	472	4	US-08-344-695-2	Sequence 17011, A
37	96.4	612	4	US-08-344-695-2	Sequence 2, Appl
38	96.4	612	4	US-08-344-695-2	Sequence 11, Appl
39	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
40	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
41	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
42	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
43	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
44	96.4	612	4	US-08-344-695-2	Sequence 14, Appl
45	96.4	612	4	US-08-344-695-2	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-428-197-48
Sequence 48, Application US/08428197
Patent No. 5891438
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,197
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-197-48
Query Match 100.0%; Score 28; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 A1SGSG 6

Db 50 A1SGSG 55

RESULT 2
PCT-US93-10555-48

Sequence 48, Application PC/TUS9310555

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.

TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATANT AND CONJUGATES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: FD-2630

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 98 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US93-10555-48

Query Match

Best Local Similarity 100.0%; Score 28; DB 5; Length 98;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 50 A1SGSG 55

RESULT 3

US-08-428-197-2

Sequence 2, Application US/08428197

Patent No. 5891438

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.

TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATANT AND CONJUGATES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/428,197

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: FD-2630

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 116 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

CLONE: ED8.4

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..116

US-08-428-197-2

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 116;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 50 A1SGSG 55

RESULT 4

US-08-983-607-36

Sequence 36, Application US/08983607

Patent No. 6140470

GENERAL INFORMATION:

APPLICANT: Alan Garen

TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-

TITLE OF INVENTION: bodies

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Department of Molecular Biophysics

STREET: 266 Whitney Avenue

CITY: New Haven

STATE: Connecticut

COUNTRY: United States of America

ZIP: 06520-8114

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" 1.44 MB diskette

COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS

SOFTWARE: Word Processing

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/983,607

FILING DATE: April 27, 1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/IB96/01032

FILING DATE: June 28, 1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens (melanoma patient immu-
ORGANISM: nized with autologous tumor cells)
INDIVIDUAL ISOLATE: peripheral blood lymphocytes
IMMEDIATE SOURCE:
LIBRARY: DM414 scfv antibodies obtained from
CLONE: V474
FEATURE:
NAME/KEY: heavy chain
US-08-983-607-36

Query Match 100.0%; Score 28; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
|||||
DB 50 A1SGSG 55

RESULT 5
PCT-US93-10555-2
Sequence 2, Application PC/TUS9310555
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATANT AND COMBUSTATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
CLONE: ED8.4
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..116
PCT-US93-10555-2

Query Match 100.0%; Score 28; DB 5; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
|||||
DB 50 A1SGSG 55

RESULT 6
US-08-545-809A-109
Sequence 109, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
US-08-545-809A-109

Query Match 100.0%; Score 28; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
|||||
DB 69 A1SGSG 74

RESULT 7
US-08-983-607-46

```
; Sequence 46, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Alan Garen
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; TITLE OF INVENTION: bodies
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Department of Molecular Biophysics
; ADDRESSEE: and Biochemistry, Yale University
; STREET: 266 Whitney Avenue
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: United States of America
; ZIP: 06520-8114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processing
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,607
; FILING DATE: April 27, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB96/01032
; FILING DATE: June 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Kinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: OCR-679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-9544
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient
; ORGANISM: immunized with autologous tumor cells)
; INDIVIDUAL ISOLATE: peripheral blood lympho-
; INDIVIDUAL ISOLATE: cytes
; IMMEDIATE SOURCE:
; LIBRARY: VH antibodies obtained from fuses
; LIBRARY: fusion phage construct
; CLONE: C55
; FEATURE:
; NAME/KEY: heavy chain
; US-08-983-607-46

Query Match          100.0%; Score 28; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 A1SGSG 6
        |||||
Db      50 A1SGSG 55
```

```
RESULT 8
; Sequence 12, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourne, JK
; APPLICANT: Allen, DJ
```

```
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; TITLE OF INVENTION: methods.
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (PRO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-652-816A-12

Query Match          100.0%; Score 28; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 A1SGSG 6
        |||||
Db      50 A1SGSG 55
```

```
RESULT 9
; Sequence 22, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Alan Garen
; APPLICANT: Xiaohong Cai
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; TITLE OF INVENTION: bodies
```

NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Department of Molecular Biophysics
ADDRESSEE: and Biochemistry, Yale University
STREET: 266 Whitney Avenue
CITY: New Haven
STATE: Connecticut
COUNTRY: United States of America
ZIP: 06520-8114
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processing
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,607
FILING DATE: April 27, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB96/01032
FILING DATE: June 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
DESCRIPTION: polypeptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens (melanoma patient immu-
INDIVIDUAL ISOLATE: peripheral blood lymphocytes
IMMEDIATE SOURCE:
LIBRARY: DM414 scfv antibodies obtained from
LIBRARY: fuses fusion phage construct
CLONE: G57
FEATURE:
NAME/KEY: heavy chain
US-08-983-607-22

Query Match 100.0%; Score 28; DB 3; Length 119;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
Db 50 AISGSG 55

RESULT 10
US-09-025-769B-38
Sequence 38, Application US/09025769B
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Illag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckthun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave

STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-769B-38

Query Match 100.0%; Score 28; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
Db 50 AISGSG 55

RESULT 11
US-09-025-769B-63
Sequence 63, Application US/09025769B
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Illag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckthun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:

NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-769B-63

Query Match 100.0%; Score 28; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
DB 50 AISGSG 55

RESULT 12
US-08-958-201-6
Sequence 6, Application US/08958201
Patent No. 5977319
GENERAL INFORMATION:
APPLICANT: Pope, Anthony R
APPLICANT: Pritchard, Kevin J
APPLICANT: Williams, Andrew J
APPLICANT: Johnson, Kevin S
TITLE OF INVENTION: Specific binding members for estradiol;
TITLE OF INVENTION: materials and methods
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/958,201
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,897
FILING DATE: 21-OCT-1996
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: 2G
US-08-958-201-6

Query Match 100.0%; Score 28; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
DB 50 AISGSG 55

RESULT 13
US-08-428-197-38

Sequence 38, Application US/08428197
Patent No. 5891438
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATANT AND CONUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jupas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,197
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-197-38

Query Match 100.0%; Score 28; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
DB 50 AISGSG 55

RESULT 14
PCT-US93-10555-38
Sequence 38, Application PC/TUS9310555
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERNATANT AND CONUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jupas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10555
; FILING DATE: 29-OCT-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-2630
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-10555-38

```

```

Query Match          100.0%; Score 28; DB 5; Length 123;
Best Local Similarity 100.0%; Pred. NO. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 A1SGSG 6
Db 50 A1SGSG 55

```

```

RESULT 15
US-08-983-607-51
; Sequence 51, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Alan Garen
; APPLICANT: Xiaohong Cai
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; TITLE OF INVENTION: bodies
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Department of Molecular Biophysics
; ADDRESSEE: and Biochemistry, Yale University
; STREET: 266 Whitney Avenue
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: United States of America
; ZIP: 06520-8114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processing
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,607
; FILING DATE: April 27, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/1996/01032
; FILING DATE: June 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Kinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: OCR-679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-9544
; TELEFAX: 203-773-1183
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient
; INDIVIDUAL ISOLATE: immunized with autologous tumor cells)
; INDIVIDUAL ISOLATE: peripheral blood lympho-
; IMMEDIATE SOURCE: cycles
; LIBRARY: VH antibodies obtained from fuses
; CLONE: E-13
; FEATURE:
; NAME/KEY: heavy chain
; US-08-983-607-51

```

```

Query Match          100.0%; Score 28; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. NO. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 A1SGSG 6
Db 50 A1SGSG 55

```

Search completed: August 20, 2003, 12:44:18
Job time : 5.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 6.50602 Seconds
(without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-27
Perfect score: 28
Sequence: 1 AISGSG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues
Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	6	11 US-09-300-425B-27	Sequence 27, Appl
2	28	100.0	17	10 US-08-192-85A-10	Sequence 10, Appl
3	28	100.0	17	10 US-09-968-561A-14	Sequence 14, Appl
4	28	100.0	17	10 US-09-968-561A-38	Sequence 38, Appl
5	28	100.0	17	10 US-09-968-561A-68	Sequence 68, Appl
6	28	100.0	17	10 US-09-968-561A-74	Sequence 74, Appl
7	28	100.0	17	10 US-09-968-561A-80	Sequence 80, Appl
8	28	100.0	17	10 US-09-968-561A-86	Sequence 86, Appl
9	28	100.0	17	10 US-09-968-561A-92	Sequence 92, Appl
10	28	100.0	17	10 US-09-968-561A-122	Sequence 122, Appl
11	28	100.0	17	10 US-09-968-561A-128	Sequence 128, Appl
12	28	100.0	17	10 US-09-968-561A-134	Sequence 134, Appl
13	28	100.0	17	10 US-09-968-561A-140	Sequence 140, Appl
14	28	100.0	17	10 US-09-968-561A-146	Sequence 146, Appl
15	28	100.0	17	10 US-09-968-561A-188	Sequence 188, Appl

16	28	100.0	17	10 US-09-968-561A-194	Sequence 194, App
17	28	100.0	17	10 US-09-968-561A-212	Sequence 212, App
18	28	100.0	17	10 US-09-968-561A-218	Sequence 218, App
19	28	100.0	17	10 US-09-968-561A-236	Sequence 236, App
20	28	100.0	17	10 US-09-968-561A-260	Sequence 260, App
21	28	100.0	17	10 US-09-968-561A-266	Sequence 266, App
22	28	100.0	17	10 US-09-968-561A-272	Sequence 272, App
23	28	100.0	17	10 US-09-968-561A-284	Sequence 284, App
24	28	100.0	17	10 US-09-968-561A-296	Sequence 296, App
25	28	100.0	17	10 US-09-968-561A-308	Sequence 308, App
26	28	100.0	17	10 US-09-968-561A-314	Sequence 314, App
27	28	100.0	17	11 US-09-972-656-45	Sequence 45, Appl
28	28	100.0	17	12 US-09-968-744A-14	Sequence 14, Appl
29	28	100.0	17	12 US-09-968-744A-38	Sequence 38, Appl
30	28	100.0	17	12 US-09-968-744A-68	Sequence 68, Appl
31	28	100.0	17	12 US-09-968-744A-74	Sequence 74, Appl
32	28	100.0	17	12 US-09-968-744A-80	Sequence 80, Appl
33	28	100.0	17	12 US-09-968-744A-86	Sequence 86, Appl
34	28	100.0	17	12 US-09-968-744A-92	Sequence 92, Appl
35	28	100.0	17	12 US-09-968-744A-122	Sequence 122, App
36	28	100.0	17	12 US-09-968-744A-128	Sequence 128, App
37	28	100.0	17	12 US-09-968-744A-134	Sequence 134, App
38	28	100.0	17	12 US-09-968-744A-140	Sequence 140, App
39	28	100.0	17	12 US-09-968-744A-146	Sequence 146, App
40	28	100.0	17	12 US-09-968-744A-188	Sequence 188, App
41	28	100.0	17	12 US-09-968-744A-194	Sequence 194, App
42	28	100.0	17	12 US-09-968-744A-212	Sequence 212, App
43	28	100.0	17	12 US-09-968-744A-218	Sequence 218, App
44	28	100.0	17	12 US-09-968-744A-236	Sequence 236, App
45	28	100.0	17	12 US-09-968-744A-260	Sequence 260, App

ALIGNMENTS

RESULT 1
US-09-300-425B-27
: Sequence 27, Application US/09300425B
: Publication No. US20030045681A1
: GENERAL INFORMATION:
: APPLICANT: NERI, Danilo
: APPLICANT: TARLI, Lorenzo
: APPLICANT: VITI, Francesca
: APPLICANT: BIRCHER, Manfred
: TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY. CONJUGATES
: TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
: TITLE OF INVENTION: ANGIOGENESIS
: FILE REFERENCE: SCH-1733P1
: CURRENT APPLICATION NUMBER: US/09/300,425B
: CURRENT FILING DATE: 1999-04-28
: PRIOR APPLICATION NUMBER: 09/075,338
: PRIOR FILING DATE: 1998-05-11
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 27
: LENGTH: 6
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
: OTHER INFORMATION: antibody clone
US-09-300-425B-27
Query Match 100.0%; Score 28; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
Db 1 AISGSG 6

RESULT 2

US-09-192-854-10
; Sequence 10, Application US/09192854
; Patent No. US20020068276A1
; GENERAL INFORMATION:
; APPLICANT: Winter, Greg
; APPLICANT: Tomlinson, Ian
; TITLE OF INVENTION: Methods for Selecting Functional Peptides
; FILE REFERENCE: 3789/72916
; CURRENT APPLICATION NUMBER: US/09/192,854
; PRIOR FILING DATE: 1998-11-17
; EARLIER APPLICATION NUMBER: 60/066,729
; PRIOR FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-192-854-10

Query Match 100.0%; Score 28; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
DB 1 AISGSG 6

RESULT 3
US-09-968-561A-14
; Sequence 14, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/10738
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-14

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
DB 1 AISGSG 6

RESULT 4
US-09-968-561A-38
; Sequence 38, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M

APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/10738
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-38

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
DB 1 AISGSG 6

RESULT 5
US-09-968-561A-68
; Sequence 68, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/10738
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-68

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
|||||
DB 1 AISGSG 6

RESULT 6
US-09-968-561A-74
; Sequence 74, Application US/09968561A

Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-74

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISSG 6
|||||
Db 1 AISSG 6

RESULT 7
US-09-968-561A-80
; Sequence 80, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-80

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISSG 6
|||||
Db 1 AISSG 6

RESULT 8
US-09-968-561A-86
; Sequence 86, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-86

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISSG 6
|||||
Db 1 AISSG 6

RESULT 9
US-09-968-561A-92
; Sequence 92, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-92

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISSG 6
|||||

Db 1 AISGSG 6

RESULT 10

US-09-968-561A-122
; Sequence 122, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968, 561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 1997-10-20
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 122
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-122

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AISGSG 6
Db 1 AISGSG 6

RESULT 11

US-09-968-561A-128
; Sequence 128, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968, 561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 1997-10-20
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 128
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-128

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
Db 1 AISGSG 6

RESULT 12

US-09-968-561A-134
; Sequence 134, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968, 561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 1997-10-20
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-134

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
Db 1 AISGSG 6

RESULT 13

US-09-968-561A-140
; Sequence 140, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968, 561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 1997-10-20
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-13
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 1998-10-20
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-968-561A-140

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SSSG 6
|||||
DB 1 A1SSSG 6

RESULT 14

US-09-968-561A-146
; Sequence 146, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:

; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/10738
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-968-561A-146

Query Match 100.0%; Score 28; DB 10; Length 17;

Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SSSG 6
|||||
DB 1 A1SSSG 6

RESULT 15

US-09-968-561A-188
; Sequence 188, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:

; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands
; FILE REFERENCE: 8039/10738
; CURRENT APPLICATION NUMBER: US/09/968,561A
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 188
; LENGTH: 17
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-968-561A-188

Query Match 100.0%; Score 28; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SSSG 6
|||||
DB 1 A1SSSG 6

Search completed: August 20, 2003, 13:16:45
Job time : 7.50602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 4.19277 Seconds
(without alignments)
137,621 Million cell updates/sec

Title: US-09-512-082-27

Perfect score: 28

Sequence: 1 AISGSG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	28	100.0	98	2 S26889	Ig heavy chain V r
2	28	100.0	99	2 S24259	Ig heavy chain V r
3	28	100.0	100	2 S24258	Ig heavy chain V r
4	28	100.0	101	2 S24257	Ig heavy chain V r
5	28	100.0	102	2 S24260	Ig heavy chain V r
6	28	100.0	104	2 S24255	Ig heavy chain V r
7	28	100.0	105	2 S24249	Ig heavy chain V r
8	28	100.0	106	2 S24256	Ig heavy chain V r
9	28	100.0	108	2 PH1648	Ig heavy chain V r
10	28	100.0	109	2 PH1649	Ig heavy chain V r
11	28	100.0	109	2 S24254	Ig heavy chain V r
12	28	100.0	109	2 S24253	Ig heavy chain V r
13	28	100.0	110	2 S24250	Ig heavy chain V r
14	28	100.0	111	2 S69911	Ig V-D-J region (R
15	28	100.0	112	2 PH1647	Ig heavy chain V r
16	28	100.0	113	2 S24247	Ig heavy chain V r
17	28	100.0	113	2 S25571	Ig heavy chain V r
18	28	100.0	117	1 H3H026	Ig heavy chain pre
19	28	100.0	117	1 A34953	Ig heavy chain pre
20	28	100.0	117	2 A34964	Ig heavy chain pre
21	28	100.0	117	2 B34964	Ig heavy chain pre
22	28	100.0	118	2 S31121	Ig heavy chain - h
23	28	100.0	119	2 C36005	Ig heavy chain V r
24	28	100.0	119	2 D36005	Ig heavy chain V r
25	28	100.0	119	2 S31107	Ig heavy chain - h
26	28	100.0	119	2 S31108	Ig heavy chain - h
27	28	100.0	120	2 S48798	Ig heavy chain V r
28	28	100.0	121	2 S31113	Ig heavy chain - h
29	28	100.0	121	2 I55673	Ig heavy chain - h

30	28	100.0	123	2 S31114	Ig heavy chain - h
31	28	100.0	125	2 S72665	Ig V-D-J region (R
32	28	100.0	127	2 S38489	Ig heavy chain - h
33	28	100.0	134	2 S31699	Ig heavy chain V r
34	28	100.0	138	2 S31666	Ig heavy chain V r
35	28	100.0	140	2 S31588	Ig heavy chain V r
36	28	100.0	140	2 S31686	Ig heavy chain V r
37	28	100.0	140	2 I47204	Ig heavy chain var
38	28	100.0	160	2 S05271	Ig heavy chain pre
39	28	100.0	182	2 F96948	probable sugar pho
40	28	100.0	183	1 C69474	conserved hypochet
41	28	100.0	186	2 F87250	molybdenum cofacto
42	28	100.0	197	1 B69131	conserved hypochet
43	28	100.0	209	2 C90155	hypothetical prote
44	28	100.0	222	2 H72582	hypothetical prote
45	28	100.0	254	1 C69073	conserved hypochet

ALIGNMENTS

RESULT 1
S26889
Ig heavy chain V region (DP-47) - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_rev1sion 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26889
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; NCID:93021117; PMID:1404388
A:Accession: S26889
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <10M>
A:Cross-references: EMBL:212347; NID:932914; PTDN:CAA78217.1; PTD:932915
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <10M>

Query Match 100.0%; Score 28; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AISGSG 6
Db 50 AISGSG 55

RESULT 2
S24259
Ig heavy chain V region (VH26-DK1-JH4) - human
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_rev1sion 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24259
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.
Submitted to the EMBL Data Library, June 1992.
A:Description: A single VH gene predominates in the rearranged and expressed human B
A:Reference number: S24247
A:Accession: S24259
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <10M>
A:Cross-references: EMBL:X67067; NID:938391; PTDN:CAA47452.1; PTD:938392
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-83/Domain: Immunoglobulin homology <10M>

Query Match 100.0%; Score 28; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AISGSG 6

Db 35 A1SGSG 40

RESULT 3

S24258

Ig heavy chain V region (VH26-DXP1-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S24258

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B cell

A:Reference number: S24247

A:Accession: S24258

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <STE>

A:Cross-references: EMBL:X67066; NID:q38389; PIDN:CAA47451.1; PID:q38390

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:4-86/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 100;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6

Db 38 A1SGSG 43

RESULT 4

S24257

Ig heavy chain V region (VH26-DXP1-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000

C:Accession: S24257

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B cell

A:Reference number: S24247

A:Accession: S24257

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-101 <STE>

A:Cross-references: EMBL:X67065; NID:q38387; PIDN:CAA47450.1; PID:q38388

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:11-93/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 101;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6

Db 45 A1SGSG 50

RESULT 5

S24260

Ig heavy chain V region (VH26-DN1-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S24260

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B cell

A:Reference number: S24247

A:Accession: S24260

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-102 <STE>

A:Cross-references: EMBL:X67068

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:5-87/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 102;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6

Db 39 A1SGSG 44

RESULT 6

S24255

Ig heavy chain V region (VH26-DLR5-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 06-Jun-1997

C:Accession: S24255

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

A:Accession: S24255

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-104 <STE>

A:Cross-references: EMBL:X67063

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:2-84/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 104;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6

Db 36 A1SGSG 41

RESULT 7

S24249

Ig heavy chain V region (VH26-DN1-DXP1-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 30-May-1997

C:Accession: S24249

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

A:Accession: S24249

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <STE>

A:Cross-references: EMBL:X67070

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:10-92/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 105;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6

Db 44 A1SGSG 49

RESULT 8

S24256

Ig heavy chain V region (VH26-DXP4-JH6) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S24256

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B cell

A:Reference number: S24247

A:Accession: S24256

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-106 <STE>

A:Cross-references: EMBL:X67064; NID:938385; PIDN:CAA47449.1; PID:938386

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:12-94/Domain: immunoglobulin homology <IMM>

Query Match

100.0%; Score 28; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6

DB 38 AISGSG 43

RESULT 9

PH1648

Ig heavy chain V region (clone 288) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996

C:Accession: PH1648

R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.

J. Exp. Med. 178, 331-336, 1993

A:Title: The structural basis of germ-line-encoded VH3 immunoglobulin binding to staphylo

A:Reference number: PH1642; MUID:93301610; PMID:8315388

A:Accession: PH1648

A:Molecule type: mRNA

A:Residues: 1-108 <HTL>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:7-90/Domain: immunoglobulin homology <IMM>

Query Match

100.0%; Score 28; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6

DB 42 AISGSG 47

RESULT 10

PH1649

Ig heavy chain V region (clone 2E7) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996

C:Accession: PH1649

R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.

J. Exp. Med. 178, 331-336, 1993

A:Title: The structural basis of germ-line-encoded VH3 immunoglobulin binding to staphylo

A:Reference number: PH1642; MUID:93301610; PMID:8315388

A:Accession: PH1649

A:Molecule type: mRNA

A:Residues: 1-109 <HTL>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:7-90/Domain: immunoglobulin homology <IMM>

Query Match

100.0%; Score 28; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6

DB 42 AISGSG 47

RESULT 11

S24254

Ig heavy chain V region (VH26-DXP2-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000

C:Accession: S24254

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

A:Accession: S24254

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-109 <STE>

A:Cross-references: EMBL:X67062

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match

100.0%; Score 28; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6

DB 49 AISGSG 54

RESULT 12

S24253

Ig heavy chain V region (VH26-DLR4-JH6) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000

C:Accession: S24253

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

A:Accession: S24253

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-109 <STE>

A:Cross-references: EMBL:X67061

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:12-94/Domain: immunoglobulin homology <IMM>

Query Match

100.0%; Score 28; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6

DB 46 AISGSG 51

RESULT 13

S24250

Ig heavy chain V region (VH26-DNL-JH4) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000

C:Accession: S24250

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

Job time : 4.19277 secs

A:Accession: S24250
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-110 <STE>
 A:Cross-references: EMBL:X67071
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 28; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 |||||
 Db 49 AISGSG 54

RESULT 14

S69911
 Ig V-D-J region (RM) - human
 C:Species: Homo sapiens (man)
 C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
 C:Accession: S69911
 R:Shiota, S.; Hamblin, T.; Oscier, D.G.; Stevenson, F.K.
 Leukemia 8, 1285-1289, 1994
 A>Title: Assessment of the role of clonogenic B lymphocytes in the pathogenesis of multi
 A:Reference number: S69909; PMID:94335315; PMID:8057663
 A:Accession: S69911
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-111 <SAH>
 A:Cross-references: EMBL:Z33401
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 28; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 |||||
 Db 50 AISGSG 55

RESULT 15

PH1647
 Ig heavy chain V region (clone 2D10) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
 C:Accession: PH1647
 R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
 J. Exp. Med. 178, 331-336, 1993
 A>Title: The structural basis of germ-line-encoded VH3 immunoglobulin binding to staphylo
 A:Reference number: PH1642; PMID:93301610; PMID:8315388
 A:Accession: PH1647
 A:Molecule type: mRNA
 A:Residues: 1-112 <HIL>
 A:Experimental source: B cell
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 28; DB 2; Length 112;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 |||||
 Db 42 AISGSG 47

Search completed: August 20, 2003, 12:42:15

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 2.27711 Seconds
(without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-27

Perfect score: 28
Sequence: 1 A1SGSG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	28	100.0	117 1 HV3C_HUMAN	P01764 homo sapien
2	28	100.0	158 1 MOAC_GAUCR	O9AC52 caulobacter
3	28	100.0	183 1 YH96_ARCFU	O28478 archaeoglob
4	28	100.0	197 1 Y249_MERTH	O26351 methanobact
5	28	100.0	222 1 Y740_AERPE	O9YAK0 aeropyrum p
6	28	100.0	414 1 SPYA_PELCA	P41689 felis silve
7	28	100.0	450 1 DHE4_LACCI	P54388 laccaria bi
8	28	100.0	457 1 DHE4_AGABI	P54387 agarticus bi
9	28	100.0	459 1 DHE4_EMENI	P18819 emericella
10	28	100.0	484 1 AM11_MYCTU	O05835 mycobacteri
11	28	100.0	1176 1 SLAP_BACSH	P38537 bacillus sp
12	27	96.4	272 1 PROC_PSEAE	P22008 pseudomonas
13	27	96.4	276 1 BACH_HALHE	O48315 halobacteri
14	27	96.4	294 1 PROC_MYCLE	P46725 mycobacteri
15	27	96.4	295 1 PROC_MYCTU	O53612 streptomyc
16	27	96.4	338 1 DHAS_STRAK	O53612 streptomyc
17	27	96.4	344 1 DHAS_CORFL	P41400 cornebacte
18	27	96.4	344 1 DHAS_CORFL	P26511 cornebacte
19	27	96.4	344 1 GUNA_TRIRE	O14405 trichoderm
20	27	96.4	345 1 DHAS_MYCBO	P47730 mycobacteri
21	27	96.4	345 1 DHAS_MYCTU	P97049 mycobacteri
22	27	96.4	346 1 DHAS_MYCSM	P41404 mycobacteri
23	27	96.4	355 1 LAVI_PHYPO	P14725 physarum po
24	27	96.4	392 1 SPYA_RABIT	P13030 oryctolagus
25	27	96.4	396 1 SPYI_CAVRO	P05995 cavia porce
26	27	96.4	430 1 ODP2_STAUV	O58821 staphylococ
27	27	96.4	446 1 DHE4_UNKP	P1657 unknown pro
28	27	96.4	447 1 DHE4_SALTI	O82616 salmonella
29	27	96.4	447 1 DHE4_SALTI	P15111 salmonella
30	27	96.4	736 1 PIRA_THEMA	O9YV22 thermotoga
31	27	96.4	931 1 POOL_ECOLI	P31828 escherichia
32	27	96.4	1850 1 BAZA_MOUSE	O91365 mus musculu
33	27	96.4	2364 1 PGCA_BOVIN	P13608 bos taurus

34	26	92.9	241 1 LEG3_RABIT	P47845 oryctolagus
35	26	92.9	249 1 LEG3_HUMAN	P17931 homo sapien
36	26	92.9	269 1 MTN_TREPA	P96122 treponema p
37	26	92.9	295 1 LEG3_CANFA	P38486 canis famli
38	26	92.9	315 1 GBI1_ARATH	P42774 arabidopsis
39	26	92.9	372 1 GDF1_HUMAN	P27539 homo sapien
40	26	92.9	373 1 L52_ADE12	P36715 human adeno
41	26	92.9	376 1 DP3B_STRPCO	P27903 streptomyc
42	26	92.9	391 1 DXR_RHIME	O92186 rhizobium m
43	26	92.9	406 1 G64B_DROME	P83284 drosophila
44	26	92.9	453 1 DHE4_NEUCR	P00359 neurospora
45	26	92.9	462 1 CH11_CANAL	P46876 candida alb

ALIGNMENTS

RESULT 1
HV3C_HUMAN STANDARD; PRT; 117 AA.

AC P01764:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region VH26 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;

RN [1]
RX MEDLINE-8110190; PubMed-6450418;
RA Mathysens G., Rabbits T.H.;
RT "Structure and multiplicity of genes for the human immunoglobulin
heavy chain variable region".
Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: J00236; AAA53516.1; -;
DR EMBL: M35415; AAA58735.1; -;
DR PIR: A02047; H3H026.
DR PDB: 1H0U; 23-DEC-99.
DR Genew: HGNC:5545; IGHV9.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig-MHC.
DR InterPro: IPR003596; Ig-V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG-LIKE; 1.
DR Immunoglobulin V region; Signal; 3D-structure.
KW SIGNAL
FT CHAIN 1 19 IG HEAVY CHAIN V-III REGION VH26.
FT DOMAIN 20 >117 IG-LIKE.
FT NON_TER 117
SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1SGSG 6
|||||||

Db 69 AISGSG 74

RESULT 2
MOAC_CAUCR STANDARD; PRT; 158 AA.

AC 09AC52; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Molybdenum cofactor biosynthesis protein C.
MOAC OR CC0014.

OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A.S., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

CC -1 FUNCTION: Together with moaA, is involved in the conversion of a
guanosine derivative (GAP) into molybdopterin precursor Z (By
similarity).
CC -1 PATHWAY: Molybdenum cofactor biosynthesis; first step.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
DR EMBL: AF005676; AKK22002.1; ALT_INT.
DR HSSP: P30747; 1EKR.
DR TIGR: CC0014; -; 1.
DR HAMAP: MF_01224; -; 1.
DR InterPro: IPR002820; Moac.
DR Pfam: PF01967; Moac; 1.
DR TIGRPFAM: TIGR00581; moac; 1.
KM Molybdenum cofactor biosynthesis; Complete proteome.
FT ACT_SITE 126 126 POTENTIAL.
SQ SEQUENCE 158 AA; 16362 MW; AC72C5B648AD7E3 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
Db 44 AISGSG 49

RESULT 3
YH96_ARCFU STANDARD; PRT; 183 AA.

AC 028478; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Hypothetical protein AFI196.
AFI196.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
RT

OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrleides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gokey J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Arlanch P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).

CC -1 SIMILARITY: BELONGS TO THE SIS FAMILY. YCKF SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
DR EMBL: AE000979; AAB89472.1; -.
DR PIR: C69474; C69474.
DR TIGR: AF1796; -.
DR InterPro: IPR001347; SIS.
DR Pfam: PF01380; SIS; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 183 AA; 20147 MW; FC2B8DA78E2EF14F CRC64;

Query Match 100.0%; Score 28; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
Db 82 AISGSG 87

RESULT 4
Y249_METHH STANDARD; PRT; 197 AA.

AC 026351; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Conserved protein.
MTH249.

OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Ducette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vitcare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";

```

RL J. Bacteriol. 179:7135-7155(1997).
CC -----
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. YCKF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
CC EMBL; AE00812; AAB84755.1; -.
CC DR PIR; B69131; B69131.
CC DR InterPro; IPR001347; SIS.
CC KW Pfam; PF01380; SIS; 1.
CC SEQUENCE 197 AA; 21570 MW; 9840D27964B6F267 CRC64;
CC
Query Match 100.0%; Score 28; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 AISGSG 6
| | | | |
Db 88 AISGSG 93
-----
RESULT 5
YV40_AERPE STANDARD; PRT; 222 AA.
ID YV40_AERPE
AC O9YAKO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein APE1940.
GN APE1940.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KI;
RX MEDLINE=99310339; PubMed=10382966;
RA Kaverbachayasi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takemiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kusuda N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Mikuchi H.;
RT *Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KI."
RL DNA Res. 6:83-101(1999).
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. YCKF SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000062; BAA80949.1; -.
CC DR PIR; H72582; H72582.
CC DR InterPro; IPR001347; SIS.
CC KW Pfam; PF01380; SIS; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 222 AA; 24339 MW; 5ED4AC8CE2B36A1 CRC64;
CC
Query Match 100.0%; Score 28; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

```

Oy      1 A1SGSG 6
        |||||
Db      111 A1SGSG 116

RESULT 6
SPYA_FELCA
ID SPYA_FELCA STANDARD; PTR; 414 AA.
AC P41689;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine--pyruvate aminotransferase, mitochondrial precursor
DE (EC 2.6.1.51) (SPT) (Alanine--glyoxylate aminotransferase)
DE (EC 2.6.1.44) (AGT).
CN AGXT OR AGT1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_Taxid=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver.
RX MEDLINE=94222101; PubMed=8168541;
RA Lamb M.J., Purdie P.E., Danpure C.J.;
RT "Molecular evolution of alanine/glyoxylate aminotransferase 1
RT Intracellular targeting. Analysis of the feline gene.";
RL Eur. J. Biochem. 221:53-62(1994).
CC -1- FUNCTION: DUAL METABOLIC ROLES OF GLUCONEOGENESIS (IN THE
CC MITOCHONDRIA) AND GLYOXYLATE DETOXIFICATION (IN THE PEROXISOMES).
CC -1- CATALYTIC ACTIVITY: L-serine + pyruvate -> 3-hydroxypyruvate + L-
CC alanine.
CC -1- CATALYTIC ACTIVITY: L-alanine + glyoxylate -> pyruvate + glycine.
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (>90%) AND PEROXISOMAL.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative Initiation;
CC Comment=2 isoforms, Mitochondrial (shown here) and Peroxisomal,
CC are produced by alternative initiation;
CC -1- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, X75923; CAAS3527.1; -.
DR PIR, S43253; S43253.
DR InterPro: IPR000192; Aminotransfv.
DR Pfam: PF00266; aminotran_5.1
DR PROSITE, PS00595; AA_TRANSFER_CLASS_5.1.
KW Transferase; Aminotransferase; Pyridoxal phosphate; Peroxisome;
KW Mitochondrion; Transit peptide; Alternative initiation.
FT TRANSIT 1 23 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 24 414 SERINE--PYRUVATE AMINOTRANSFERASE,
FT ISOREM MITOCHONDRIAL,
FT CHAIN 23 414 SERINE--PYRUVATE AMINOTRANSFERASE,
FT ISOREM PEROXISOMAL.
FT INIT MET 23 23 FOR ISOREM PEROXISOMAL.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT SITE 412 414 MICROBODY TARGETING SIGNAL (BY
FT SIMILARITY)
SQ SEQUENCE 414 AA; 45507 MW; 0D1B01E0E9A199B3 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AISGSG 6
Db      99 AISGSG 104

RESULT 7
DHE4_LACBI STANDARD; PRT; 450 AA.
ID DHE4_LACBI
AC P54388;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE NADP-specific glutamate dehydrogenase (EC 1.4.1.4) (NADP-GDH).
GN GDH.
OS Laccaria bicolor.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Tricholomataceae; Laccaria.
OX NCBI_TaxID=29883;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S238N;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -1 SUBUNIT: Homohexamer (By similarity).
CC -1 SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U13369; AAA82936.1; -
DR HSSP: P24295; 1AUP.
DR InterPro: IPR006095; GLFY_dehydrog.
DR InterPro: IPR006096; GLFY_dehydrog_C.
DR InterPro: IPR006097; GLFY_dehydrog_N.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF00208; GLFY_dehydrog_1.
DR Pfam: PF02812; GLFY_dehydrog_N_1.
DR PRINTS: PR00082; GLFYDHGRNASE.
DR PROSITE: PS00074; GLFY_DEHYDROGENASE; 1.
DR OXidoreductase; NADP.
DR ACT SITE 111
DR FT ACT SITE 111 BY SIMILARITY.
SQ SEQUENCE 450 AA; 48474 MW; B803A30FBFAF22090 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 450;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AISGSG 6
Db      223 AISGSG 228

RESULT 8
DHE4_AGABI STANDARD; PRT; 457 AA.
ID DHE4_AGABI
AC P54387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE NADP-specific glutamate dehydrogenase (EC 1.4.1.4) (NADP-GDH).
GN GDH.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Horst H39;
RA MEDLINE=96180651; PubMed=8602149;
RA Schaepe P.J., Mueller Y., Baars J.J.P., Op den Camp H.J.M.,
RA Sonnenberg A.S.M., van Griensven L.J.L.D., Visser J.;
RT "Nucleotide sequence and expression of the gene encoding NADP+-
RT dependent glutamate dehydrogenase (gdh) from Agaricus bisporus.";
RL Mol. Gen. Genet. 250:339-347(1996).
CC -1 CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -1 SUBUNIT: Homohexamer (By similarity).
CC -1 SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X83393; CA58312.1; -
DR PTR: S63608; S63608.
DR HSSP: P24295; 1AUP.
DR InterPro: IPR006095; GLFY_dehydrog.
DR InterPro: IPR006096; GLFY_dehydrog_C.
DR InterPro: IPR006097; GLFY_dehydrog_N.
DR Pfam: PF00208; GLFY_dehydrog_1.
DR Pfam: PF02812; GLFY_dehydrog_N_1.
DR PRINTS: PR00082; GLFYDHGRNASE.
DR PROSITE: PS00074; GLFY_DEHYDROGENASE; 1.
DR OXidoreductase; NADP.
DR ACT SITE 111
DR FT ACT SITE 111 BY SIMILARITY.
SQ SEQUENCE 457 AA; 49557 MW; 1BF0E97F67078AC4 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AISGSG 6
Db      223 AISGSG 228

RESULT 9
DHE4_EMENT STANDARD; PRT; 459 AA.
ID DHE4_EMENT
AC P18619;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE NADP-specific glutamate dehydrogenase (EC 1.4.1.4) (NADP-GDH).
GN GDH.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89384423; PubMed=2550758;
RA Hawkins A.R., Guir S.J., Montague P., Kinghorn J.R.;
RT "Nucleotide sequence and regulation of expression of the Aspergillus
RT nidulans gdhA gene encoding NADP dependent glutamate dehydrogenase.";
RL Mol. Gen. Genet. 218:105-111(1989).
CC -1 CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -1 SUBUNIT: Homohexamer.
CC -1 SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: X16121; CA34252.1; -
 DR HSSP: P24295; 1AUP.
 DR InterPro: IPR006095; GLFV_dehydrog.
 DR InterPro: IPR006095; GLFV_dehydrog_C.
 DR InterPro: IPR006097; GLFV_dehydrog_N.
 DR Pfam: PF02812; GLFV_dehydrog_1.
 DR Pfam: PF02812; GLFV_dehydrog_N; 1.
 DR PRINTS: PR00082; GLFV_DEHYDROGENASE.
 DR PROSITE: PS00074; GLFV_DEHYDROGENASE; 1.
 DR OXidoreductase; NADP.
 KM ACT_SITE 114 114 BY SIMILARITY.
 SQ SEQUENCE 459 AA; 49608 MW; 682964399C00287D CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AISGSG 6
 Db 225 AISGSG 230
 RESULT 10
 AM1_MYCTU STANDARD; PRT; 484 AA.
 ID AM1_MYCTU
 AC O05835;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative amidase am1a2 (EC 3.5.1.4).
 GN AM1A2 OR RV2363 OR MT2432 OR MTC127.17C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Baccocck K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Felkell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikulac A.,
 RA Bissel W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: A monocarboxylic acid amide + H(2)O = a
 CC -1- monocarboxylate + NH(3).
 CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: Z95208; CAB08460.1; -
 DR EMBL: AE007082; AAK46726.1; -
 DR PIR: B70586; B70586.
 DR TIGR: MT2432; -
 DR TubercuList: Ry2363; -
 DR InterPro: IPR00120; Amidase.
 DR Pfam: PF01425; Amidase; 1.
 DR PROSITE: PS00571; AMIDASES; 1.
 DR Hypothetical protein; Hydrolase; Complete proteome.
 KM SEQUENCE 484 AA; 50883 MW; 90D86CCD0C90F02A CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 484;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AISGSG 6
 Db 11 AISGSG 16
 RESULT 11
 SLAP_BACSH STANDARD; PRT; 1176 AA.
 ID SLAP_BACSH
 AC P38537;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Surface-layer 125 kDa protein precursor.
 OS Bacillus sphaericus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC NCBI_TaxID=1421;
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=2362;
 RX MEDLINE=89327128; PubMed=2666389;
 RA Bowditch R.D., Baumann P., Yousten A.A.;
 RA "Cloning and sequencing of the gene encoding a 125-kilodalton
 RT surface-layer protein from Bacillus sphaericus 2362 and of a related
 RT cryptic gene."
 RL J. Bacteriol. 171:4178-4188(1989).
 CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
 CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
 CC -1- SUBCELLULAR LOCATION: Cell wall.
 CC -1- SIMILARITY: Contains 3 S-layer homology (SLH) domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: M28361; AAA50256.1; -
 DR PIR: A33856; A33856.
 DR InterPro: IPR001119; SLH.
 DR Pfam: PF00395; SLH; 3.
 DR PROSITE: PS01072; SLH_DOMAIN; 2.
 KM Signal; Cell wall; S-layer; Repeat.
 RT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 1 1176 SURFACE-LAYER 125 kDa PROTEIN.
 FT DOMAIN 32 91 SLH 1.
 FT DOMAIN 92 151 SLH 2.
 FT DOMAIN 152 210 SLH 3.
 SQ SEQUENCE 1176 AA; 125225 MW; 1A9A458BF433788C CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 1176;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 DB 746 AISGSG 751

RESULT 12

PROC_PSEAE STANDARD; PRT; 272 AA.
 ID AC P22008;
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
 GN PROC OR PA0393.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
 RC STRAIN-ATCC 15692 / PA01;
 RA MEDLINE=90185238; PubMed=2107123;
 RA Savioz A., Jeunes D.J., Kocher H.P., Haas D.;
 RT "Comparison of proc and other housekeeping genes of Pseudomonas
 RT aeruginosa with their counterparts in Escherichia coli.";
 RL Gene 86:107-111(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RA MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lapidis K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-6 FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RA MEDLINE=91285432; PubMed=1676385;
 RA Whitchurch C.B., Hobbs M., Livingston S.P., Krishnapillai V.,
 RA Mettick J.S.;
 RT "Characterisation of a Pseudomonas aeruginosa twitching motility gene
 RT and evidence for a specialised protein export system widespread in
 RT eubacteria.";
 RL Gene 101:33-44(1991).
 CC -1- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
 CC carboxylate + NAD(P)H.
 CC -1- PATHWAY: Proline biosynthesis; third (last) step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M3357; AAA25975.1;
 CC EMBL: AE004476; AAG03782.1;
 CC EMBL: M55524; AAA25958.1;
 CC PIR: J00418; J00418.
 CC InterPro: IPR00304; P5CR.
 CC Pfam: PF01089; P5CR.1.
 CC TIGRfam: TIGR00112; PROC.1.
 CC PROSITE: PS00521; P5CR.1.

KM Oxidoreductase: Proline biosynthesis; NADP: Complete proteome.
 FT INIT MET 0
 SQ SEQUENCE 272 AA; 27963 MW; 750520E6028D73E0 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 272;
 Best Local Similarity 83.3%; Pred. No. 99;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 DB 170 AMSGSG 175

RESULT 13

BACH_HALHP STANDARD; PRT; 276 AA.
 ID AC Q48315;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Halorhodopsin precursor (HR).
 GN HOP.
 OS Halobacterium halobium (strain port).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=33004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96130249; PubMed=8541296;
 RA Odomo J., Muramatsu T.;
 RT "Over-expression of a new photo-active halorhodopsin in Halobacterium
 RT salinarum.";
 RL Biochim. Biophys. Acta 1240:248-256(1995).
 CC -1- FUNCTION: HALORHODOPSIN IS A LIGHT-DRIVEN ANION PUMP. MAXIMAL
 CC ABSORPTION IS 578 NM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL OPSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: D43766; BAA07823.1;
 CC PIR: T48843; T48843.
 CC HSP: P16102; IE12.
 CC InterPro: IPR001425; Bac_rhodopsin.
 CC Pfam: PF01036; Bac_rhodopsin; 1.
 CC PRINTS: PR00251; BACTERIOPSIN.
 CC PROSITE: PS00327; BACTERIAL_OPSIN_REF. 1.
 CC PROSITE: PS00950; BACTERIAL_OPSIN_1; 1.
 CC Ion transport; Photoreceptor; Transmembrane; Retinal protein.
 CC BY SIMILARITY.
 CC CHAIN 1 21
 CC DOMAIN 22 276
 CC TRANSMEM 28 27
 CC TRANSMEM 63 51
 CC TRANSMEM 104 122
 CC TRANSMEM 133 156
 CC TRANSMEM 160 184
 CC TRANSMEM 193 216
 CC TRANSMEM 228 251
 CC DOMAIN 252 276
 CC BINDING 241 241
 CC SEQUENCE 276 AA; 29045 MW; A2D024F474E59307 CRC64;
 Query Match 96.4%; Score 27; DB 1; Length 276;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AISGSG 6
|:||||
Db 258 AVS GSG 263

RESULT 14

PROC_MYCLE STANDARD; PRT; 294 AA.
ID PROC_MYCLE
AC PA6725;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyrolytic-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PROC OR ML2430 OR B2168_C2.211.
OS Mycobacterium lepre.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE-21128732; PubMed-11234002;
RA Cole S.T., Elgimeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Hough R.S., Hornsby T., Jagels K., Lacroix C., Mclean J., Moulé S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) -> 1-pyrraline-5-
CC carboxylate + NAD(P)H.
CC -1- PATHWAY: Proline biosynthesis: third (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: U00018; AAA1233.1; -
CC EMBL: AL883925; CAC31947.1; -
CC PIR: S72897; S72897.
CC Leproma; ML2430; -
CC InterPro: IPR000304; P5CR.
CC Pfam: PF01089; P5CR.1.
CC TIGRFAMs: TIGR00112; PROC.1.
CC PROSITE: PS00521; P5CR.1.
CC Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
KW SEQUENCE 294 AA; 30237 MW; EA8606C9CBB6BD9D CRC64;
SQ

Query Match 96.4%; Score 27; DB 1; Length 294;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AISGSG 6
|:||||
Db 176 AVS GSG 181

RESULT 15

PROC_MYCTU STANDARD; PRT; 295 AA.
ID PROC_MYCTU
AC 01141;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyrolytic-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PROC OR RV0500 OR MT0520 OR MTCY20C9.26.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Elgimeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Hough R.S.,
RA Hornsby T., Jagels K., Krogh A., Mclean J., Moulé S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Kolonay A., Utterback T., Weidman J., Khouli H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) -> 1-pyrraline-5-
CC carboxylate + NAD(P)H.
CC -1- PATHWAY: Proline biosynthesis: third (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: Z77162; CAB00926.1; -
CC EMBL: AE006952; AAK4743.1; -
CC PIR: G70745; G70745.
CC TIGR: MT0520; -
CC Tuberculist: RV0500; -
CC InterPro: IPR000304; P5CR.
CC Pfam: PF01089; P5CR.1.
CC TIGRFAMs: TIGR00112; PROC.1.
CC PROSITE: PS00521; P5CR.1.
CC Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
KW SEQUENCE 295 AA; 30171 MW; F95C3F407BE5408F CRC64;
SQ

Query Match 96.4%; Score 27; DB 1; Length 295;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AISGSG 6
|:||||
Db 177 AVS GSG 182

Wed Aug 20 13:35:36 2003

us-09-512-082-27.rsp

Page 8

Search completed: August 20, 2003, 12:34:52
Job time : 4.38822 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 11.3494 Seconds
(Without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-27
Perfect score: 28
Sequence: 1 AISGSG 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	182	16 097M03	097M03 clostridium
2	28	100.0	195	17 097G52	097G52 sulfobolus
3	28	100.0	202	17 082X34	082X34 pyrobaculum
4	28	100.0	209	17 0980X4	0980X4 sulfobolus
5	28	100.0	210	5 09NJV7	09NJV7 cryptospori
6	28	100.0	216	17 08TR01	08TR01 methanosarc
7	28	100.0	219	17 08PUG7	08PUG7 methanosarc
8	28	100.0	254	17 027589	027589 methanobact
9	28	100.0	271	5 09N8A7	09N8A7 trypanosoma
10	28	100.0	274	16 08XVZ1	08XVZ1 ralsstonia s
11	28	100.0	291	16 08NCH7	08NCH7 corynebacte
12	28	100.0	359	16 097LJ2	097LJ2 clostridium
13	28	100.0	364	2 093TV0	093TV0 pseudomonas
14	28	100.0	370	3 094129	094129 aspergillus
15	28	100.0	379	5 081192	081192 drosophila
16	28	100.0	392	11 035504	035504 cavia porce

17	28	100.0	406	17 059601	059601 pyrococcus
18	28	100.0	408	17 09UX25	09UX25 pyrococcus
19	28	100.0	418	3 09HG52	09HG52 debaryomyce
20	28	100.0	426	16 09K9H6	09K9H6 bacillus ha
21	28	100.0	434	17 08U472	08U472 pyrococcus
22	28	100.0	445	2 0923C4	0923C4 pseudomonas
23	28	100.0	445	16 09HVJ7	09HVJ7 pseudomonas
24	28	100.0	450	3 093934	093934 botrytis ci
25	28	100.0	450	3 096U99	096U99 hebeloma cy
26	28	100.0	451	3 097804	097804 schizosacch
27	28	100.0	457	3 09HPR6	09HPR6 tuber borch
28	28	100.0	461	3 09UR51	09UR51 penicillium
29	28	100.0	476	16 08YU75	08YU75 anabaena sp
30	28	100.0	551	16 08YP28	08YP28 anabaena sp
31	28	100.0	552	5 09VE00	09VE00 drosophila
32	28	100.0	562	10 08GRU7	08GRU7 oryza sativ
33	28	100.0	597	4 096BB9	096BB9 homo sapien
34	28	100.0	691	5 09VW11	09VW11 drosophila
35	28	100.0	739	5 096680	096680 drosophila
36	28	100.0	739	16 08XWV1	08XWV1 ralsstonia s
37	28	100.0	830	12 083863	083863 nllaparvata
38	28	100.0	843	9 034076	034076 streptococc
39	28	100.0	900	2 08GH57	08GH57 pseudomonas
40	28	100.0	1025	16 08PJ02	08PJ02 xanthomonas
41	28	100.0	1208	16 09PF19	09PF19 xyella fas
42	28	100.0	1227	5 095Z44	095Z44 leishmania
43	28	100.0	1636	16 09PE17	09PE17 xyella fas
44	28	100.0	2023	5 09V529	09V529 drosophila
45	28	100.0	2023	5 096342	096342 drosophila

ALIGNMENTS

RESULT 1

ID Q97M03 PRELIMINARY; PRT; 182 AA.
AC Q97M03;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Predicted sugar phosphate aminotransferase involved in capsule formation.
GN CAC0397.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.W., Dubois J., Qiu D., Hiltl J., Wolf Y.I., Ratusev R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007554; AAK78377.1; -;
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
KW Transferrase; Aminotransferase; Complete proteome.
SQ SEQUENCE 182 AA: 19724 MW: 5F0E612061ACF588 CRC64;

Query Match 100.0%; Score 28; DB 16; Length 182;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
DB 81 AISGSG 86

RESULT 2
 0976J2 PRELIMINARY; PRT; 195 AA.
 AC 0976J2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein ST0198.
 GN ST0198.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Tanaka Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagi M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.;
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AF000981; BAB5155.1; -
 DR InterPro: IPR001347; SIS.
 DR Pfam: PF01380; SIS; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 195 AA; 21350 MW; ABAD918947DB8BF CRC64;

Query Match 100.0%; Score 28; DB 17; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 Db 84 AISGSG 89

RESULT 3
 082X34 PRELIMINARY; PRT; 202 AA.
 AC 082X34;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Conserved protein with sugar isomerase (SIS) domain.
 GN PAE1489.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AE009824; AAL6315.1; -
 DR InterPro: IPR001347; SIS.
 DR Pfam: PF01380; SIS; 1.
 KW Isomerase; Complete proteome.
 SQ SEQUENCE 202 AA; 22092 MW; 138068682FE1C45A CRC64;

Query Match 100.0%; Score 28; DB 17; Length 202;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 Db 90 AISGSG 95

RESULT 4
 0980X4 PRELIMINARY; PRT; 209 AA.
 AC 0980X4;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE D-arabino-3-hexulose 6-phosphate formaldehyde lyase (hps-1).
 GN HPS-1 OR SS00151.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: AE006654; AAK0498.1; -
 DR InterPro: IPR001347; SIS.
 DR Pfam: PF01380; SIS; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 209 AA; 22980 MW; 6DE48593FDEAE114 CRC64;

Query Match 100.0%; Score 28; DB 17; Length 209;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AISGSG 6
 Db 98 AISGSG 103

RESULT 5
 09NJV7 PRELIMINARY; PRT; 210 AA.
 AC 09NJV7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Proteasome B type subunit.
 GN PB.
 OS Cryptosporidium parvum.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 OC Cryptosporididae; Cryptosporidium.
 OX NCBI_TaxID=5807;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KSU-1;
 RX MEDLINE=20541292; PubMed=11092745;
 RA Chung P.A., Johnson J., Khramtsov N.V., Upton S.J.;
 RT "Cloning and molecular characterization of a gene encoding a
 RT Cryptosporidium parvum putative 20S proteasome beta-type subunit.";
 RL DNA Seq. 11:309-314(2000).
 CC -1- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
 CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG,
 CC PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR
 CC SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC
 CC ACTIVITY (BY SIMILARITY).

```

CC -1- PATHWAY: INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (BY SIMILARITY).
CC EMBL: AF097521; AAF72737.1; -.
DR HSSP: P38624; 1RYP.
DR InterPro: IPR000243; Proteasome_B.
DR InterPro: IPR001353; Proteasome_Protease.
DR Pfam: PF00227; Proteasome; 1.
DR PRINTS: PR00141; PROTEASOME.
DR PROSITE: PS00854; PROTEASOME_B; 1.
KW Hydrolyase; Protease; Proteasome.
SQ SEQUENCE 210 AA; 22962 MW; 1F1352F3540692C7 CRC64;

Query Match 100.0%; Score 28; DB 5; Length 210;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AISGSG 6
Db 139 AISGSG 144

RESULT 6
Q8TR01 PRELIMINARY; PRT; 216 AA.
ID Q8TR01;
AC Q8TR01;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 6-phospho-3-hexulolomerase.
GN MA1384.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbam C., Roy A., Endrizzi M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atroor D., Brown A.,
RA Allen N., Maylor J., Stange-Thomann N., DeRellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talmas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-547(2002).
DR EMBL: AE010808; AM04800.1; -.
DR InterPro: IPR001347; SIS.
DR Pfam: PF01380; SIS; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 216 AA; 23538 MW; 1237FBA5E824FECC3 CRC64;

Query Match 100.0%; Score 28; DB 17; Length 216;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AISGSG 6
Db 105 AISGSG 110

RESULT 7

```

```

Q8PUG7 PRELIMINARY; PRT; 219 AA.
ID Q8PUG7;
AC Q8PUG7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hexulose-6-phosphate isomerase (EC 5.3.1.1).
GN NM2367.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gael / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wlezer A., Baumer S., Jacobl C.,
RA Brueggemann H., Llenard T., Christmann A., Boemecke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013479; AM32063.1; -.
DR InterPro: IPR001347; SIS.
DR Pfam: PF01380; SIS; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 219 AA; 24122 MW; 9A08818BCD2C8103 CRC64;

Query Match 100.0%; Score 28; DB 17; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AISGSG 6
Db 108 AISGSG 113

RESULT 8
Q27589 PRELIMINARY; PRT; 254 AA.
ID Q27589;
AC Q27589;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Conserved protein.
GN MTH1546.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeRougey C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pochier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gidson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AE000915; AAB86020.1; -.
DR InterPro: IPR001347; SIS.
DR Pfam: PF01380; SIS; 1.
KW Complete proteome.
SQ SEQUENCE 254 AA; 28218 MW; 3671BD3FF8D674A8 CRC64;

Query Match 100.0%; Score 28; DB 17; Length 254;

```

Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
Db 144 A1SGSG 149

RESULT 9

O9N8A7 PRELIMINARY; PRT; 271 AA.
AC O9N8A7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Hypothetical 30.5 kDa protein.
GN CHL1.424.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TREU927;
RA Hall N., Bowman S., Quail M., Ivans A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: A1929607; CAB95638.1;
KW Hypothetical protein.
SQ SEQUENCE 271 AA; 30459 MW; C821823052478090 CRC64;

Query Match 100.0%; Score 28; DB 5; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
Db 225 A1SGSG 230

RESULT 10

O8XVZ1 PRELIMINARY; PRT; 274 AA.
AC O8XVZ1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Probable oxidoreductase pyrroline-5-carboxylate reductase signal
DE peptide protein (EC 1.5.1.2).
GN PROC OR RSC2684 OR RS00037.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EM11000;
RX MEDLINE=21618179; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cartolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Layte M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siglier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).
DR EMBL: AL646071; CAD16391.1;
DR InterPro: IPR000304; PSOR.
DR Pfam: PF01089; PSOR; 1.
DR TIGRfams: TIGR00112; PROC; 1.
DR PROSITE: PS00521; PSOR; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 274 AA; 28113 MW; C0CD1P9546003BEB CRC64;

Query Match 100.0%; Score 28; DB 16; Length 274;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
Db 171 A1SGSG 176

RESULT 11

O8NOH7 PRELIMINARY; PRT; 291 AA.
AC O8NOH7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Transcriptional regulators.
GN CGL1457.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RL "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005278; BAB98850.1;
DR InterPro: IPR000281; HTH_RPir.
DR InterPro: IPR001347; SIS.
DR Pfam: PF01418; HTH_6; 1.
KW Complete proteome.
SQ SEQUENCE 291 AA; 30257 MW; 31AC86B3F4A0C25E CRC64;

Query Match 100.0%; Score 28; DB 16; Length 291;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A1SGSG 6
Db 191 A1SGSG 196

RESULT 12

O97LJ2 PRELIMINARY; PRT; 359 AA.
AC O97LJ2;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Aspartate semialdehyde dehydrogenase (Gene asd).
GN CAC0568.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007572; AAK78547.1;
DR InterPro: IPR000534; Semialdh_dh.
DR Pfam: PF01118; Semialdh_dh; 1.

DR Pfam: PF02774; Semialdehyde_dhc; 1.
 KW Complete proteome.
 SQ SEQUENCE 359 AA; 40383 MW; A6E336B1F05C6065 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 16; Length 359;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 185 AISGSG 190

RESULT 13

ID O93TV0 PRELIMINARY; PRT; 364 AA.
 AC O93TV0;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glutamate dehydrogenase GdhA (Fragment).
 GN GdhA.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM1;
 RA Syn C.K.C., Magnuson J., Kingsley M.T., Swarup S.;
 RT "Pseudomonas putida glutamate dehydrogenase (gdhA), partial cds."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF311093; AAK49517.2; -
 DR HSSP: P96110; 1826.
 DR InterPro: IPR006095; GLFV_dehydrog.
 DR InterPro: IPR006096; GLFV_dehydrog_C.
 DR InterPro: IPR006097; GLFV_dehydrog_N.
 DR Pfam: PF00208; GLFV_dehydrog; 1.
 DR Pfam: PF02812; GLFV_dehydrog_N; 1.
 DR PRINTS: PR00082; GLFDHRCNASE.
 DR PROSITE: PS0074; GLFV_DEHYDROGENASE; 1.
 FT NON_TER 1
 SQ SEQUENCE 364 AA; 38520 MW; 0384628DA36D2448 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 2; Length 364;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 153 AISGSG 158

RESULT 14

ID O94129 PRELIMINARY; PRT; 370 AA.
 AC O94129;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE NADP-dependent glutamate dehydrogenase.
 GN GDHA.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC Eucotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillius.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 22342;
 RA Cardoza R.E., Moralejo F.J., Gutierrez S., Casquelro J., Fierro F.,
 RT "Characterization and nitrogen source regulation at the
 RT transcriptional level of the gdhA gene of Aspergillus awamori encoding

RT an NADP-dependent glutamate dehydrogenase.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y15784; CAA75779.1; -
 DR HSSP: P24295; 1AUP.
 DR InterPro: IPR006095; GLFV_dehydrog.
 DR InterPro: IPR006096; GLFV_dehydrog_C.
 DR InterPro: IPR006097; GLFV_dehydrog_N.
 DR Pfam: PF00208; GLFV_dehydrog; 1.
 DR Pfam: PF02812; GLFV_dehydrog_N; 1.
 DR PRINTS: PR00082; GLFDHRCNASE.
 DR NADP.
 KW SEQUENCE 370 AA; 40103 MW; AE2D8BA32F8E08E9 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 3; Length 370;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 135 AISGSG 140

RESULT 15

ID O81192 PRELIMINARY; PRT; 379 AA.
 AC O81192;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE RH2-PA.
 GN RH2.
 OS Drosophila littoralis.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=47316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tucson 15010-1001.10;
 RX MEDLINE-22426072; PubMed-12537575;
 RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A.,
 RA Guirke A., Mungall C.J., Wang A.M., Kronmiller B., Pacleb J., Park S.,
 RA Stapleton M., Wan K., George R.A., de Jong P., Bots J., Rubin G.M.,
 RA Cealiker S.E.;
 RT "Assessing the impact of comparative genomic sequence data on the
 RT functional annotation of the Drosophila genome."
 RL Genome Biol. 3:RESEARCH0086.1-RESEARCH0086.20(2002).
 DR EMBL: AY190955; AA001080.1; -
 SQ SEQUENCE 379 AA; 42266 MW; 07B6ADE9C97AE5EC6 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 5; Length 379;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AISGSG 6
 |||||
 DB 8 AISGSG 13

Search completed: August 20, 2003, 12:40:18
 Job time : 14.4605 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 12:35:21 ; Search time 7.56627 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-29
Perfect score: 18
Sequence: 1 GLSI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	18	100.0	10	6 Q9TR48	Q9TR48 bos taurus
2	18	100.0	14	4 Q960P2	Q960P2 homo sapien
3	18	100.0	31	16 Q9PGY3	Q9PGY3 xylella tas
4	18	100.0	34	5 Q23184	Q23184 caenorhabd
5	18	100.0	34	5 Q23303	Q23303 caenorhabd
6	18	100.0	34	5 Q23306	Q23306 caenorhabd
7	18	100.0	34	5 Q17647	Q17647 caenorhabd
8	18	100.0	35	8 Q62956	Q62956 picea abies
9	18	100.0	35	15 Q76284	Q76284 human immun
10	18	100.0	35	15 Q76283	Q76283 human immun
11	18	100.0	35	15 Q76282	Q76282 human immun
12	18	100.0	37	16 Q8KAT7	Q8KAT7 chlorobium
13	18	100.0	42	2 P74973	P74973 xanthomonas
14	18	100.0	45	3 Q43098	Q43098 gibberella
15	18	100.0	46	2 Q04528	Q04528 xanthomonas
16	18	100.0	48	16 Q8EJ74	Q8EJ74 shewanella

17	18	100.0	48	16 Q8E515	Q8E515 streptococc
18	18	100.0	49	2 Q53860	Q53860 spiroplasma
19	18	100.0	49	12 Q88421	Q88421 spiroplasma
20	18	100.0	49	16 Q8CLV9	Q8CLV9 yersinia pe
21	18	100.0	50	10 Q8GXB2	Q8GXB2 arabidopsis
22	18	100.0	53	16 Q8YKB6	Q8YKB6 mycobacteri
23	18	100.0	55	16 Q8XST3	Q8XST3 anabaena sp
24	18	100.0	56	3 Q8X180	Q8X180 emericelela
25	18	100.0	56	8 Q8E8C7	Q8E8C7 guillardi
26	18	100.0	56	16 Q8E8C7	Q8E8C7 streptococ
27	18	100.0	56	16 Q8E8C7	Q8E8C7 streptococ
28	18	100.0	56	16 Q8E8C7	Q8E8C7 streptococ
29	18	100.0	58	12 Q8E8C7	Q8E8C7 hepatitis b
30	18	100.0	58	16 Q8E8C7	Q8E8C7 hepatitis b
31	18	100.0	58	16 Q8E8C7	Q8E8C7 hepatitis b
32	18	100.0	59	16 Q8E8C7	Q8E8C7 hepatitis b
33	18	100.0	59	16 Q8E8C7	Q8E8C7 hepatitis b
34	18	100.0	61	2 Q48651	Q48651 lactococcus
35	18	100.0	61	6 Q95L65	Q95L65 felis silve
36	18	100.0	62	16 Q95L65	Q95L65 felis silve
37	18	100.0	62	16 Q95L65	Q95L65 felis silve
38	18	100.0	63	16 Q8E8C7	Q8E8C7 hepatitis b
39	18	100.0	64	11 Q9CT63	Q9CT63 mus muscu
40	18	100.0	64	16 Q8X1Q0	Q8X1Q0 raietonia s
41	18	100.0	64	16 Q8D5H2	Q8D5H2 vibrio vuln
42	18	100.0	65	16 Q9JYS1	Q9JYS1 neisseria m
43	18	100.0	65	16 Q984E4	Q984E4 rhizobium l
44	18	100.0	65	16 Q99YK2	Q99YK2 streptococ
45	18	100.0	65	16 Q8D740	Q8D740 streptococ

ALIGNMENTS

RESULT 1

Q9TR48 ID Q9TR48 PRELIMINARY; PRT; 10 AA.
AC Q9TR48; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Amphoterin (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96029671; PubMed=7592757;
RA Hori O., Brett J., Slatery T., Cao R., Zhang J., Chen J.X.,
RA Nagashima M., Lundh E.R., Vijay S., Nitecki D.;
RT "The receptor for advanced glycation end products (RAGE) is a cellular
RT binding site for amphoterin. Mediation of neurite outgrowth and co-
RT expression of RAGE and amphoterin in the developing nervous system.";
RT J. Biol. Chem. 270:25752-25761(1995).
RL J. Biol. Chem. 270:25752-25761(1995).
SQ SEQUENCE 10 AA; 1050 MW; 23B89A26761EB18 CRC64;

Query Match 100.0%; Score 18; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
7 GLSI 10

RESULT 2
Q960P2 ID Q960P2 PRELIMINARY; PRT; 14 AA.
AC Q960P2; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

175

DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE Tissue factor pathway inhibitor-2 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hube F., Reverdiou-Moalic P., Iochmann S., Cherpi-Antar C., Gruel Y.,
 RT "Characterization of the human tissue factor pathway inhibitor-2
 (TFPI-2) gene promoter region."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 FR EMBL, AY044097; AAK72693.1;
 DT NON_TER 14
 SQ SEQUENCE 14 AA; 1509 MW; 4E70ED601BC1177 CRC64;
 Query Match 100.0%; Score 18; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 DB 8 GLSI 11
 RESULT 3
 O9PGY3 PRELIMINARY; PRT; 31 AA.
 AC O9PGY3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DE 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DE 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 GN Hypothetical protein Xf0163.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5c;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
 RA Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Facinanci A.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Fraga J.S., Ferreira A.J.S., Ferreira V.C.A., Ferris J.A.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Klieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukawa A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
 RA de Souza A.P., Terenzi M.F., Truifi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Setubal J.C.,
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159(2000).
 DR EMBL, AEO03870; AAF82976.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 31 AA; 3348 MW; 51A14263E9F91A3 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 31;
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 DB 10 GLSI 13
 RESULT 4
 O23184 PRELIMINARY; PRT; 34 AA.
 AC O23184;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DE 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 DE W06A7.5 protein.
 GN W06A7.5
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ainscough R.,
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99065613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL, Z78066; CAB01519.1;
 DR WormPep: W06A7.5; CE14672.
 SQ SEQUENCE 34 AA; 3601 MW; 659917D9DE1DA528 CRC64;
 Query Match 100.0%; Score 18; DB 5; Length 34;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 DB 23 GLSI 26
 RESULT 5
 O23303 PRELIMINARY; PRT; 34 AA.
 AC O23303;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DE 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 DE ZC412.7 protein.
 GN ZC412.7 OR H12D21.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ainscough R.,
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA McMurtry A.,
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL, Z78067; CAB01526.1;
 DR EMBL, Z92849; CAB07428.1;
 DR WormPep: H12D21.1; CE15223.
 DR WormPep: ZC412.7; CE15223.
 SQ SEQUENCE 34 AA; 3619 MW; 6598A7D9DE1DA528 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 34;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 23 GLSI 26

RESULT 6
 023306 ID 023306 PRELIMINARY; PRT; 34 AA.
 AC 023306; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE ZC412.6 protein.
 GN ZC412.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Almscough R.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RT Science 282:2012-2018(1998).
 DR EMBL; Z78067; CAB01529.1; -.
 DR WormPeP; ZC412.6; CE15221.
 SQ SEQUENCE 34 AA; 3653 MW; 6C08A7D9DE1DA528 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 34;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 23 GLSI 26

RESULT 7
 017647 ID 017647 PRELIMINARY; PRT; 34 AA.
 AC 017647; 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE C40H5.1 protein.
 GN C40H5.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA White S.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RT Science 282:2012-2018(1998).
 DR EMBL; Z61462; CAB03952.1; -.
 DR WormPeP; C40H5.1; CE086649.

SQ SEQUENCE 34 AA; 3599 MW; 0EE8A7DB894DA529 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 34;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 23 GLSI 26

RESULT 8
 062956 ID 062956 PRELIMINARY; PRT; 35 AA.
 AC 062956; 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Orf35.
 OS Picea abies (Norway spruce) (Picea excelsa).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 OX NCBI_TaxID=3329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kluemper S., Kanka S., Riesner D., Etscheid M.;
 RT "Characterisation of a Norway spruce chloroplast DNA clone: complete
 nucleotide sequences of rpl23, rpl2, rps19, rpl22, rps3, trn I,
 pseudo-ndhc and a residual inverted repeat B.";
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U92462; AAC95502.1; -.
 SQ SEQUENCE 35 AA; 3922 MW; 4D1E72F51281B5B7 CRC64;

Query Match 100.0%; Score 18; DB 8; Length 35;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 18 GLSI 21

RESULT 9
 076284 ID 076284 PRELIMINARY; PRT; 35 AA.
 AC 076284; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HIV-1[GUN-4];
 RA Shimizu N.S.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HIV-1[GUN-4];
 RX MEDLINE=94335139; PubMed=7980782;
 RA Shimizu N.S., Shimizu N.G., Takeuchi Y., Hoshino H.;
 RT "Isolation and characterization of human immunodeficiency virus type 1
 variants Infectious to brain-derived cells: detection of common point
 mutations in the V3 region of the env gene of the variants.";
 RT J. Virol. 68:6130-6135(1994).
 DR EMBL; D34603; BAA07002.1; -.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120.1.

```

KM AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
RC STRAIN-HIV-1[GUN-4];
SQ SEQUENCE 35 AA; 3847 MW; A3670DE7FA40F415 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
   ||||
Db 11 GLSI 14

RESULT 10
Q76283 PRELIMINARY; PRT; 35 AA.
ID Q76283
AC Q76283;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Envelope glycoprotein (Fragment).
ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HIV-1[GUN-4];
RA Shimizu N.S.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-HIV-1[GUN-4];
RA MEDLINE=94335139; PubMed=7980782;
RX Shimizu N.S., Shimizu N.G., Takeuchi Y., Hoshino H.;
RT "Isolation and characterization of human immunodeficiency virus type 1
RT variants Infectious to brain-derived cells: detection of common point
RT mutations in the V3 region of the env gene of the variants.";
RL J. Virol. 68:6130-6135(1994).
DR EMBL; D34602; BAA07001.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
SQ SEQUENCE 35 AA; 3819 MW; A3670DE57FA40F415 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
   ||||
Db 11 GLSI 14

RESULT 11
Q76282 PRELIMINARY; PRT; 35 AA.
ID Q76282
AC Q76282;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Envelope glycoprotein (Fragment).
ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HIV-1[GUN-4];
RA Shimizu N.S.;

```

```

RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-HIV-1[GUN-4];
RX MEDLINE=94335139; PubMed=7980782;
RA Shimizu N.S., Shimizu N.G., Takeuchi Y., Hoshino H.;
RT "Isolation and characterization of human immunodeficiency virus type 1
RT variants Infectious to brain-derived cells: detection of common point
RT mutations in the V3 region of the env gene of the variants.";
RL J. Virol. 68:6130-6135(1994).
DR EMBL; D34601; BAA07000.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
SQ SEQUENCE 35 AA; 3819 MW; A3670DE57FA40F415 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
   ||||
Db 11 GLSI 14

RESULT 12
Q8KAT7 PRELIMINARY; PRT; 37 AA.
ID Q8KAT7
AC Q8KAT7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE Hypothetical protein CT2068.
GN CT2068.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE012956; AAM73285.1; -.
DR TIGR; CT2068; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 37 AA; 4349 MW; 545A837202D7F219 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 16; Length 37;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
   ||||
Db 3 GLSI 6

RESULT 13
P74973 PRELIMINARY; PRT; 42 AA.
ID P74973
AC P74973;
DT 01-FEB-1997 (TREMBLrel. 02, Created)

```

```

DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Putative histidine kinase (Fragment).
OS Xanthomonas campestris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BXPF65.
RA Chan J.W.Y.F., Maynard S., Goodwin P.H.;
RT "A two-component signal transduction system of Xanthomonas campestris
RT pv. phaseoli var. fuscans strain BXPF65."
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U69111; AAB09063.1; -.
DR InterPro: IPR003594; AtPbind_ATPase.
DR InterPro: IPR005467; His_kinase.
DR Pfam: PF02518; HATPase_c; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
KW Kinase.
FT NON_TER
SQ SEQUENCE 42 AA; 4487 MW; 5B3FD26C4DE212BC CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 5 GLSI 8

RESULT 14
ID 043098 PRELIMINARY; PRT; 45 AA.
AC 043098;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Galactose oxidase (Fragment).
GN GAOA.
OS Glberella zeae (Fusarium graminearum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Nectriaceae; Glberella.
OX NCBI_TaxID=5518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 4527;
RA Nielsen M.L., Vogel R.F.;
RT "Specific identification of Fusarium graminearum by PCR with gaoA
RT targeted primers."
RL Syst. Appl. Microbiol. 20:111-123(1997).
DR EMBL; U51094; AAB94635.1; -.
DR HSSP; Q01745; 1GOF.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000421; FAS5_C.
DR PROSITE: PS00018; EF_HAND_1.
FT NON_TER
SQ SEQUENCE 45 AA; 4927 MW; 1AA403C08523D564 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 3; Length 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 11 GLSI 14

RESULT 15
ID 004528 PRELIMINARY; PRT; 46 AA.

```

```

AC 004528;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Genes for sensor and regulator protein (Fragment).
OS Xanthomonas campestris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-pathovar: campestris;
RX MEDLINE=91042416; Pubmed=2233675;
RA Osbourn A.E., Clarke B.R., Stevens B.J.H., Daniels M.J.;
RT "Use of oligonucleotide probes to identify members of two-component
RT systems in Xanthomonas campestris pathovar campestris."
RL Mol. Gen. Genet. 222:145-151(1990).
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSOR TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL; X54015; CA37962.1; -.
DR InterPro: IPR003594; AtPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR005467; His_kinase.
DR Pfam: PF02518; HATPase_c; 1.
DR PRINTS; PR00344; BCTRSSENSOR.
DR PROSITE: PS50109; HIS_KIN; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
FT NON_TER
SQ SEQUENCE 46 AA; 4824 MW; 1A3C266837B0F6C5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 46;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 9 GLSI 12

Search completed: August 20, 2003, 12:40:22
Job time: 11.6774 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 : Search time 8.79518 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: us-09-512-082-29

Perfect score: 18
Sequence: 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 45 summaries

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	8	AAU98275	M. tuberculosis Ag
2	18	100.0	9	AAK51217	A generic Bradykin
3	18	100.0	9	AAK82116	Melanoma-specific
4	18	100.0	10	AAK51218	A generic Bradykin
5	18	100.0	10	AAK95974	Human complement
6	18	100.0	10	AAK86718	Saccharomyces cere
7	18	100.0	10	AAK86810	Saccharomyces cere
8	18	100.0	10	AAK86812	Saccharomyces cere
9	18	100.0	14	AAK67162	Bovine glial cell

10	18	100.0	14	AAK86638	GGF I tryptic pept
11	18	100.0	15	AAK49865	Sequence of tryptic
12	18	100.0	15	AAK18171	Immunodominant epi
13	18	100.0	15	AAK75598	M. tuberculosis 30
14	18	100.0	15	AAK75599	M. tuberculosis 30
15	18	100.0	15	AAK65578	Oestrogen receptor
16	18	100.0	15	AAK65578	Human cytochrome P450
17	18	100.0	15	AAK86751	Human cytochrome P450
18	18	100.0	15	AAK86752	Human cytochrome P450
19	18	100.0	15	AAK86753	Human cytochrome P450
20	18	100.0	15	AAK98126	Interferon-gamma s
21	18	100.0	15	AAK98127	Interferon-gamma s
22	18	100.0	15	AAK97736	gp100 derived IFN-
23	18	100.0	15	AAK97737	gp100 derived IFN-
24	18	100.0	15	AAK103068	Hepatitis C virus
25	18	100.0	15	AAK866315	Oestrogen receptor
26	18	100.0	15	AAK56362	Mycobacterium tube
27	18	100.0	15	AAK56363	Mycobacterium tube
28	18	100.0	16	AAK67158	Bovine glial cell
29	18	100.0	16	AAK67160	Bovine glial cell
30	18	100.0	16	AAK86634	GGF I tryptic pept
31	18	100.0	16	AAK86636	Mycobacterium tube
32	18	100.0	16	AAK12281	Mycobacterium tube
33	18	100.0	16	AAK98655	Human cell death p
34	18	100.0	16	AAK46175	Human Blys binding
35	18	100.0	16	AAK46175	Human Blys binding
36	18	100.0	16	AAK46175	Human Blys binding
37	18	100.0	16	AAK11945	Human HGPBMY11 N-
38	18	100.0	16	AAK11957	Human HGPBMY11 N-
39	18	100.0	16	AAK11958	Human HGPBMY11 N-
40	18	100.0	17	AAK10034	Neutralising domai
41	18	100.0	17	AAK83375	G protein-coupled
42	18	100.0	19	AAK50695	Helicobacter pylori
43	18	100.0	20	AAK30406	Human retinoid aci
44	18	100.0	21	AAK39064	Gp100 melanoma ant
45	18	100.0	21	AAK39070	Gp100 melanoma ant

ALIGNMENTS

RESULT 1	AAU98275	standard; peptide: 8 AA.
ID	AAU98275	
AC	AAU98275	
XX		
DT	15-AUG-2002	(first entry)
XX		
DE	M. tuberculosis Ag85 derived immunodominant T cell epitope #2.	
XX		
KW	Ag85; epitope; vaccine; tuberculosis; mycolyltransferase; HLA-A*02001;	
KW	Mycobacterium tuberculosis; Mycobacterium leprae; Mycobacterium bovis;	
KW	Mycobacterium ulcerans; Mycobacterium avium; antibacterial;	
KW	MHC class I-restricted CD8+ T cell; human leukocyte antigen.	
XX		
OS	Mycobacterium tuberculosis.	
XX		
FT	Key	Location/Qualifiers
FT	Region	2
FT		/note- "Amino acid contributing to the HLA-A*02001
FT		binding motif"
FT	Region	8
FT		/note- "Amino acid contributing to the HLA-A*02001
FT		binding motif"
EP	EP1211260-A1.	
PN	05-JUN-2002.	
XX		
PD	30-NOV-2000; 2000EP-0204268.	
XX		
PF		
XX	30-NOV-2000; 2000EP-0204268.	
PR		

XX (ZIEK-) ACAD ZIEKENHUIS LEIDEN.
 PA Ottenhof THM, Geluk A;
 XX
 XX
 PI
 XX
 DR WPI: 2002-464923/50.
 XX

PT Peptide of 8-11 amino acids derived from the Ag85 protein of
 PT Mycobacterium induce proliferation of MHC class I-restricted CD8+ T
 PT cells and are useful to vaccinate against infection by mycobacterium
 XX
 PS Claim 1; Page 6; 20pp; English.

CC The invention relates to peptide (P1) derived from an Ag85 protein of
 CC Mycobacterium (which is associated with mycolyltransferase activity
 CC and is involved in cell wall synthesis), comprising 8-11 amino acids, and
 CC capable of inducing proliferation of MHC class I-restricted CD8+ T cells
 CC in vivo. Also included are (1) a peptide comprising P1 flanked by amino
 CC acids representing antigen processing sites; (2) a polypeptide comprising
 CC at least two of the peptides in (1); (3) a nucleic acid encoding one of
 CC the claimed peptides or polypeptides; (4) a vector comprising the above
 CC nucleic acid; (5) a host cell comprising the above nucleic acid or
 CC vector; and (6) detecting and/or enumerating CD8+ T cells against
 CC Mycobacterium, comprising tetrameric complexes of MHC (major
 CC histocompatibility group) class I and one of the claimed peptides or
 CC polypeptides. The molecules of the invention are used to prepare a
 CC vaccine against mycobacterium (e.g. Mycobacterium tuberculosis,
 CC Mycobacterium leprae, Mycobacterium bovis, Mycobacterium ulcerans,
 CC Mycobacterium avium) and thus protect against diseases such as
 CC tuberculosis. The present sequence is an M. tuberculosis Ag85
 CC derived immunogenic peptide of the invention binding to HLA-A*02001
 CC (human leukocyte antigen).
 CC
 SQ Sequence 8 AA;

QY Query Match 100.0%; Score 18; DB 23; Length 8;
 DB Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
 DB 1 GLSI 4

RESULT 2
 AAR51217
 ID AAR51217 standard; peptide; 9 AA.
 XX
 AC AAR51217;
 XX

DT 25-MAR-2003 (updated)
 DT 02-NOV-1994 (first entry)
 XX

DE A generic Bradykinin antagonist peptide.

KM Bradykinin antagonist peptide; 5-position; non-aromatic residue
 KW Treatment of inflammatory reactions; reduce pain.
 XX
 XX Synthetic.

OS
 XX
 FH Key Location/Qualifiers
 FT Modified-site 3 /label= 4HYP
 FT /note= "trans"
 FT Misc-difference 7 /note= "D-form residue"

PN WO9406453-A1.

XX 31-MAR-1994.

PF 08-SEP-1993; 93WO-US08220.
 XX

PR 11-SEP-1992; 92US-0942317.

XX (GERA/) GERA L.
 PA (SRIV/) SRIVASTAVA V.
 PA (STEW/) STEWART J M.
 XX

PI Gera L, Srivastava V, Stewart JM;

XX WPI: 1994-118152/14.

PT Antagonistic bradykinin analogues with non-aromatic amino acid in
 PT the 5-position - are useful for treating inflammatory reactions,
 PT hypotension, etc.
 XX

PS Example 44; Page 19; 29 Pages; English.

CC AAR51217 is an example of a generic bradykinin antagonist peptide
 CC having a non-aromatic residue in the 5-position. The antagonist
 CC peptide inhibits the bradykinin response when injected as a bolus
 CC admixture of bradykinin plus antagonist by either the i.v. or i.v.
 CC route of administration, or when administered as an infusion.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 9 AA;

QY Query Match 100.0%; Score 18; DB 15; Length 9;
 DB Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
 DB 4 GLSI 7

RESULT 3
 AAR82116
 ID AAR82116 standard; peptide; 9 AA.
 XX

AC AAR82116;

DT 25-MAR-1996 (first entry)
 XX

DE Melanoma-specific mutant immunogen epitope 9mer peptide.

KM Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;
 KW cytotoxic T cell; lymphocyte; HLA-A2.
 XX

OS Homo sapiens.

PN WO9522561-A2.

PD 24-AUG-1995.

PF 16-FEB-1995; 95WO-US01991.

PR 29-APR-1994; 94US-0234784.

PR 16-FEB-1994; 94US-0197399.

PA (UYVI-) UNIV VIRGINIA PATENT FOUND.

PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

DR WPI: 1995-302688/39.

PT Melanoma-specific immunogen comprises epitope(s) homologous with
 PT pMEL.17 - are highly potent stimulators of HLA-A2+CTL's useful in
 PT adoptive immuno-therapy
 XX
 XX
 PS Example 8; Page 51; 148pp; English.

CC A melanoma-specific immunogen homologous with pMEL-17 comprises one
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group AAR82098-
 AAR82194 capable of eliciting a CTL response. The epitopes AAR82098-

CC AAR82108 are of particular interest. The immunogen can be used for
 CC partial protection in mammals against melanoma peptides which are
 CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+
 CC CTLs in several cell lines and can be used in immunotherapy or
 CC incorporated into immunogenic conjugates as vaccines.

XX Sequence 9 AA;

Query Match 100.0%; Score 18; DB 16; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLST 4
 ||||
 DB 1 GLST 4

RESULT 4

AAR51218 standard; peptide; 10 AA.

XX AAR51218;

XX 25-MAR-2003 (updated)
 DT 02-NOV-1994 (first entry)

XX A generic Bradykinin antagonist peptide.

XX Bradykinin antagonist peptide; 5-position; non-aromatic residue
 KW Treatment of inflammatory reactions; reduce pain.

XX Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 1
 FT Modified-site 4 /note= "D-form residue"
 FT /label= "4Hyp"
 FT /note= "trans"
 FT Misc-difference 8 /note= "D-form residue"
 FT

XX WO9406453-A1.

XX 31-MAR-1994.

XX 08-SEP-1993; 93WO-US08220.

XX 11-SEP-1992; 92US-0942317.

XX (GERA/) GERA L.

XX (SRIV/) SRIVASTAVA V.

XX (STEW/) STEWART J M.

XX Gera L, Srivastava V, Stewart JM;

XX WPI; 1994-118152/14.

XX Antagonistic bradykinin analogues with non-aromatic amino acid in
 PT the 5-position - are useful for treating inflammatory reactions,
 PT hypotension, etc.

XX Example 45; Page 19; 29 Pages; English.

CC AAR51217 is an example of a generic bradykinin antagonist peptide
 CC having a non-aromatic residue in the 5-position. The antagonist
 CC peptide inhibits the bradykinin response when injected as a bolus
 CC admixture of bradykinin plus antagonist by either the ia. or iv.
 CC route of administration, or when administered as an infusion.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 10 AA;

Query Match 100.0%; Score 18; DB 15; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLST 4
 ||||
 DB 5 GLST 8

RESULT 5

AAG95974 standard; peptide; 10 AA.

XX AAG95974;

XX 18-SEP-2001 (first entry)

XX Human complementary peptide, SEQ ID NO: 2168.

XX Human; complementary peptide; ligand; drug discovery; drug design.

XX Homo sapiens.

XX WO200142277-A2.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000MO-GB04776.

XX 13-DEC-1999; 99GB-0029464.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-408419/43.

XX A set of peptide ligands consisting of specific complementary peptides
 PT to proteins encoded by genes of the human genome, useful in an assay
 PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs -

XX Example 4; Page 353; 646pp; English.

XX The invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA;

Query Match 100.0%; Score 18; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLST 4
 ||||
 DB 6 GLST 9

RESULT 6

AAG86718 standard; peptide; 10 AA.

XX AAG86718;

XX 11-SEP-2001 (first entry)

XX Saccharomyces cerevisiae peptide, SEQ ID NO: 1667.

XX Saccharomyces cerevisiae; complementary peptide; peptide identification;

```

KW drug discovery; drug design.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04773.
XX
PR 13-DEC-1999; 99GB-0029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-367863/38.
XX
PT

Example 3; Page 260; 488pp; English.

PT Identifying complementary peptides by analysis of protein and
PT nucleotide sequence databases, useful in drug design -
XX
PS
XX
CC The invention relates to the identification of complementary peptides
CC by analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents
CC and drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 18; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 GLST 4
   ||||
Db 5 GLSI 8

RESULT 7
AAG86810
ID AAG86810 standard; Peptide; 10 AA.
XX
AC AAG86810;
XX
DT 11-SEP-2001 (first entry)
XX
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1759.
XX
KW Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04773.
XX
PR 13-DEC-1999; 99GB-0029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-367863/38.
XX
PT Identifying complementary peptides by analysis of protein and

```

```

PT nucleotide sequence databases, useful in drug design -
XX
PS Example 3; Page 272; 488pp; English.
XX
CC The invention relates to the identification of complementary peptides
CC by analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents
CC and drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae.
XX
SQ Sequence 10 AA;
XX
Query Match 100.0%; Score 18; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 GLSI 4
Db 4 GLSI 7
XX
RESULT 8
AAG86812
ID AAG86812 standard; Peptide; 10 AA.
XX
AC AAG86812;
XX
DT 11-SEP-2001 (first entry)
XX
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1761.
XX
XX Saccharomyces cerevisiae; complementary peptide; peptide identification;
XX drug discovery; drug design.
XX
KM Saccharomyces cerevisiae.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04773.
XX
PR 13-DEC-1999; 99GB-0029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
WPI; 2001-367863/38.
XX
PT Identifying complementary peptides by analysis of protein and
PT nucleotide sequence databases, useful in drug design -
XX
XX Example 3; Page 272; 488pp; English.
XX
CC The invention relates to the identification of complementary peptides
CC by analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents
CC and drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae.
XX
SQ Sequence 10 AA;
XX
Query Match 100.0%; Score 18; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GLSI 4
 DB 3 GLSI 6

RESULT 9
 AAR67162
 ID AAR67162 standard; peptide; 14 AA.

AC AAR67162;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-AUG-1995 (first entry)

XX Bovine glial cell growth factor-1 (GGF-1) peptide fragment.
 DE
 XX Glial cell growth factor-1; GGF-1; mammalian muscle cell treatment;
 KW skeletal; cardiac; smooth; acetylcholine receptor deficiency.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note="Lys, Arg"
 FT
 XX W09426298-A1.
 XX
 XX 24-NOV-1994.
 XX
 PF 06-MAY-1994; 94WO-US05083.
 XX
 PR 06-MAY-1993; 93US-0059022.
 PR 08-MAR-1994; 94US-0209204.
 XX
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE.
 XX
 PI Gwynne DI, Marchionni M, Sklar R;
 XX
 DR WPI: 1995-006353/01.
 XX
 PT Treating mammalian muscle diseases and disorders - by admin. of
 PT GGF2 and other specified polypeptide(s) which bind the p185erbB2
 PT receptor.
 XX
 PS Example 9; Page 90; 241pp; English.
 XX
 CC AAR67153-R67163 and AAR67174-R67183 are bovine glial cell growth
 CC factor-1 (GGF-1) peptide fragments. A 30-36kd polypeptide factor
 CC known to have glial cell mitogenic activity, which includes one
 CC of the above peptide fragments in its amino acid sequence is
 CC claimed. The polypeptide can be used to treat a variety of
 CC mammalian skeletal, cardiac and smooth muscle diseases,
 CC including acetylcholine receptor deficiency.
 CC (Updated on 25-MAR-2003 to correct FN field.)
 CC
 XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 18; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 DB 6 GLSI 9

RESULT 10
 AAR86638
 ID AAR86638 standard; peptide; 14 AA.
 XX
 AC AAR86638;
 XX
 DT 03-JUL-1996 (first entry)

XX GGF I tryptic peptide GGF-I 10.
 DE
 XX Glial growth factor; GGF; human; hGFR2; Schwann cell; mitogenesis; GGF-I;
 KW glial cell; therapy; peripheral nerve damage; demyelination; bovine;
 KW glial tumour; fibroblast proliferation; wound repair; multiple sclerosis;
 KW neurodegenerative disorder; neural regeneration; acetylcholine receptor.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /label= Lys, Arg
 FT
 XX W09532724-A1.
 XX
 XX 07-DEC-1995.
 XX
 PF 25-MAY-1995; 95WO-US06846.
 XX
 PR 26-MAY-1994; 94US-0249322.
 XX
 PA (CAMP-) CAMBRIDGE NEUROSCIENCE.
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Chen MS, Goodearl AD, Hiles I, Marchionni M, Minnett L;
 PI Stroobant P, Waterfield M;
 XX
 DR WPI: 1996-030329/03.
 XX
 XX Glial growth factors with Schwann cell mitogenic activity - for
 PT prophylaxis or treatment of nerve disorders, e.g. Multiple sclerosis
 PT
 XX Example 2; Fig 9; 199pp; English.
 PS
 XX AAR86629-R66657 represent fragments of bovine glial growth factor I
 CC (GGF-I). These fragments were obtained by lysyl endopeptidase and
 CC protease V8 digests. These sequences have Schwann cell mitogenic
 CC activity in the presence of foetal calf plasma. These sequences (and
 CC human GGF2, see AAR86628) are used to stimulate mitogenesis of glial
 CC cells, for prophylaxis or treatment of a pathological condition of
 CC the nervous system in a mammal. Also, for identification of a receptor,
 CC for treatment of conditions of peripheral nerve damage (e.g.
 CC demyelination/damage/loss of Schwann cells), treatment of
 CC neurodegenerative disorders in mammals, for inducing neural
 CC regeneration, fibroblast proliferation or wound repair. Glial tumours
 CC can be treated by inhibiting this sequence binding to its receptor.
 CC These peptides are also able to induce synthesis of acetylcholine
 CC receptor, and is useful in prophylaxis or treatment of multiple
 CC sclerosis in a patient.
 CC
 XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 18; DB 17; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 DB 6 GLSI 9

RESULT 11
 AAR49865
 ID AAR49865 standard; peptide; 15 AA.
 XX
 AC AAR49865;
 XX
 DT 25-MAR-2003 (updated)
 DT 12-SEP-1994 (first entry)

DE Sequence of tryptic digest peptide of bovine glial growth factor III
 (GGF III).

```

XX  Glial growth factor; GGF III; mitogen; Schwann cell.
KW  Bos taurus.
XX
OS  WO9404560-A1.
XX
PN  03-MAR-1994.
XX
PD  13-AUG-1993; 93WO-GB01721.
XX
PF  14-AUG-1992; 92GB-0017316.
XX
PR  (LUDW-) LUDWIG INST CANCER RES.
XX
PA  Goodearl ADJ, Stroobant F, Waterfield MD, Goodearl ADJ;
XX  WPI, 1994-083104/10.
XX
DR  New polypeptide factor and peptide(s) from bovine pituitary -
XX  having mitogenic activity in stimulating division of Schwann
XX  cells, used for therapy, prophylaxis, diagnosis of
XX  neuro-degenerative disease, glial cell tumours, etc
XX
PS  Claim 32; Page 31; 44pp; English.
XX
CC  A novel polypeptide was purified from bovine pituitaries. It has
XX  mitogenic activity stimulating the division of Schwann cells, and
XX  exhibits a mol. wt. of 43-35KD when carrying native glycosylation. It
XX  was digested with trypsin and lysylendopeptidase to obtain novel
XX  peptides AAR49858-R49866 and AAR49867-R49871 respectively. When peptides
XX  AAR49862-R49866 were sequenced to completion it was found that none of
XX  these sequences is apparently related to GGF-I or GGF-II peptide
XX  sequences. A polypeptide contg. any of the sequences in AAR49858-
XX  AAR49871 is claimed, as is DNA encoding each of the peptides.
XX  (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ  Sequence 15 AA;
XX
Query Match 100.0%; Score 18; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSI 4
DB 7 GSI 10

```

RESULT 12

```

AAW18171
ID AAW18171 standard; peptide; 15 AA.
XX
AC AAW18171;
XX
DT 13-AUG-1997 (first entry)
XX
DE Immunodominant epitope from Mycobacterium tuberculosis 30KD protein.
XX
KW Vaccine; vaccinating agent; M.tuberculosis; pathogen; bacteria;
XX  virus; fungus; protozoan; HIV.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9637219-A1.
XX
PD 28-NOV-1996.
XX
PF 23-MAY-1996; 96WO-US07781.
XX
PR 06-DEC-1995; 95US-0566357.
XX  23-MAY-1995; 95US-0447398.
XX  20-OCT-1995; 95US-0545926.
XX  31-OCT-1995; 95US-0551149.
XX
PA

```

```

XX  (REGC ) UNIV CALIFORNIA.
XX
PA Harth G, Horwitz MA;
XX
DR WPI, 1997-020936/02.
XX
PT Vaccines derived from M.tuberculosis major abundant extracellular
XX  proteins - are easy to prepare and less toxic than conventional
XX  killed or attenuated vaccines, useful for protecting against or
XX  treating Mycobacterial infections
XX
PS Claim 45; Page 161; 193pp; English.
XX
CC A vaccinating agent for promoting an immune response in a mammal
XX  against Mycobacterium pathogens comprises at least one
XX  immunodominant epitope of at least one majorly abundant
XX  extracellular protein, i.e. the M.tuberculosis 110, 80, 71, 58, 45,
XX  32A, 32B, 30, 24, 23.5, 23, 16, 14 or 12 kD proteins, or their
XX  analogues, homologues and subunits. The present sequence represents an
XX  immunodominant epitope from the major abundant extracellular product
XX  30 kD protein. The vaccinating agents are used to protect against (or
XX  to treat existing) infections by Mycobacterium (especially M.
XX  tuberculosis) while the epitopes can also be used to detect presence
XX  of an immune response to a Mycobacterium pathogen. The vectors,
XX  containing the DNA for the extracellular proteins, are used to transform
XX  cells for production of recombinant DNA molecules. More generally the
XX  DNA from other pathogens can be used in vaccines, e.g. against other
XX  bacteria, viruses, fungi and protozoa. Since different combinations of
XX  DNA can be used, a wide range of effective compositions can be produced.
XX  They generate a response against the antigens most often found on
XX  infected cells during the infection, regardless of the strength or
XX  specificity of the immune response. The vaccines are easy to produce
XX  and less toxic than known killed or attenuated vaccines, so can be given
XX  to immunocompromised subjects, e.g. those with HIV infection.
XX
SQ Sequence 15 AA;
XX
Query Match 100.0%; Score 18; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSI 4
DB 5 GSI 8

```

RESULT 13

```

AAW75598
ID AAW75598 standard; peptide; 15 AA.
XX
AC AAW75598;
XX
DT 23-OCT-1998 (first entry)
XX
DE M. tuberculosis 30 kD protein derived peptide 12 (residues 56-70).
XX
KW Mycobacterium tuberculosis; vaccination; extracellular product;
XX  immunodominant epitope; interleukin-12; MFS9; immune response;
XX  opsonising humoral response; intracellular pathogen.
XX
OS Synthetic.
XX
PN WO9831388-A1.
XX
PD 23-JUL-1998.
XX
PF 15-JAN-1998; 98WO-US00942.
XX
PR 21-JAN-1997; 97US-0786533.
XX
PA (REGC ) UNIV CALIFORNIA.

```

XX
PI Harth G, Horwitz MA, Lee B;
XX
DR MPI: 1998-413815/35.
XX
PT Vaccines against Mycobacterium containing major extracellular
PT proteins - used to, e.g. induce protective and therapeutic immune
PT responses, and for detecting an immune response
PS
XX Example 28: Page 96; 236pp; English.
XX
CC Sequences shown in AAW75587 to AAW75641 represent synthetic peptides
CC derived from the native 30kD major secretory protein of M. tuberculosis.
CC These peptides are used for splenic lymphocyte proliferation assays to
CC identify the immunodominant T-cell epitope of the 30kD protein. The
CC invention provides an agent for vaccinating mammals against
CC Mycobacterium. The agent comprises at least one of the major abundant
CC extracellular 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or
CC 12 kD proteins of M. tuberculosis, or at least 1 of their immunodominant
CC epitopes and Interleukin-12 (IL-12) or MF59 as adjuvants. The agent
CC containing the nucleic acid encoding the extracellular products are used
CC to raise a protective or therapeutic immune response against
CC Mycobacterium, specifically M. tuberculosis. The immunodominant epitopes
CC can also be used (typically in a cutaneous hypersensitivity test) to
CC detect an immune response to vaccination. Preparation of the agent does
CC not require selection of the most immunogenic products, so large scale
CC production and purification are easy, resulting in a consistent,
CC standardised formulation, having lower toxicity than killed or attenuated
CC vaccines. The agents provide a rapid and effective response (including a
CC strong cell-mediated component) and are safe even in immunocompromised
CC subjects. They prevent development of an opsonising humoral response that
CC might spread intracellular pathogens.
XX
SQ Sequence 15 AA:
XX
Query Match 100.0%; Score 18; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
DB 10 GLSI 13

RESULT 14
AAW75599
ID AAW75599 standard; peptide: 15 AA.
XX
AC AAW75599;
XX
DE 23-OCT-1998 (first entry)
XX
DE M. tuberculosis 30 kD protein derived peptide 13 (residues 61-75).
XX
XX Mycobacterium tuberculosis; vaccination; extracellular product;
KW immunodominant epitope; interleukin-12; MF59; immune response;
KW opsonising humoral response; intracellular pathogen.
XX
OS Synthetic.
OS Mycobacterium tuberculosis.
XX
PN WO9831388-A1.
XX
PD 23-JUL-1998.
XX
PF 15-JAN-1998; 98WO-US00942.
XX
PR 21-JAN-1997; 97US-0786533.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Harth G, Horwitz MA, Lee B;
XX

DR MPI: 1998-413815/35.
XX
XX
PT Vaccines against Mycobacterium containing major extracellular
PT proteins - used to, e.g. induce protective and therapeutic immune
PT responses, and for detecting an immune response
PS
XX Example 28: Page 96; 236pp; English.
XX
CC Sequences shown in AAW75587 to AAW75641 represent synthetic peptides
CC derived from the native 30kD major secretory protein of M. tuberculosis.
CC These peptides are used for splenic lymphocyte proliferation assays to
CC identify the immunodominant T-cell epitope of the 30kD protein. The
CC invention provides an agent for vaccinating mammals against
CC Mycobacterium. The agent comprises at least one of the major abundant
CC extracellular 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or
CC 12 kD proteins of M. tuberculosis, or at least 1 of their immunodominant
CC epitopes and Interleukin-12 (IL-12) or MF59 as adjuvants. The agent
CC containing the nucleic acid encoding the extracellular products are used
CC to raise a protective or therapeutic immune response against
CC Mycobacterium, specifically M. tuberculosis. The immunodominant epitopes
CC can also be used (typically in a cutaneous hypersensitivity test) to
CC detect an immune response to vaccination. Preparation of the agent does
CC not require selection of the most immunogenic products, so large scale
CC production and purification are easy, resulting in a consistent,
CC standardised formulation, having lower toxicity than killed or attenuated
CC vaccines. The agents provide a rapid and effective response (including a
CC strong cell-mediated component) and are safe even in immunocompromised
CC subjects. They prevent development of an opsonising humoral response that
CC might spread intracellular pathogens.
XX
SQ Sequence 15 AA:
XX
Query Match 100.0%; Score 18; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
DB 5 GLSI 8

RESULT 15
AAW65578
ID AAW65578 standard; peptide: 15 AA.
XX
AC AAW65578;
XX
DE 01-FEB-2000 (first entry)
XX
DE Oestrogen receptor beta ERE binding peptide 17B-beta.
XX
XX Oestrogen receptor; estrogen; estradiol; oestrogen response element;
KW ERE; binding; biological activity; fingerprint; molecular braille;
KW cellular braille; modulation; tamoxifen; breast cancer; ovarian cancer;
KW menopause; osteoporosis; selective oestrogen receptor modulator;
KW identification; characterisation; classification.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9954728-A2.
XX
PD 28-OCT-1999.
XX
PF 26-MAR-1999; 99WO-US006664.
XX
PR 23-APR-1998; 98US-0082756.
XX
PR 09-SEP-1998; 98US-009656.
XX
PR 08-JAN-1999; 99US-0115345.
XX
PA (NOVA-) NOVALON PHARM CORP.
XX
PI Paige LA, Hamilton PT, Fowlkes DM, Buehrer B, Barnett T;
XX

PI McDonnell DP, Christensen DJ;
 XX
 DR WPI; 2000-013281/01.
 XX

PT Methods for identifying new receptor modulators, especially estrogen
 XX modulators to treat tamoxifen refractory breast cancer -
 XX

PS Example 2.2; Page 162; 219pp; English.
 XX

CC The present invention describes a method for predicting the biological
 CC activity of new receptor modulating compounds (1) using novel oligomeric
 CC peptides (bikeys) which have differential abilities to bind to 2
 CC different receptor conformations. The method is used to identify new
 CC drugs that are physiological or pharmacological agonists/antagonists and
 CC that target various receptors, which are involved in certain disease
 CC conditions. The system may be used as a primary screening tool to
 CC identify hits, to classify lead compounds from a drug screen to,
 CC characterise selective oestrogen receptor modulators (SERMs) in terms of
 CC agonist and antagonist function and to predict possible clinical effects
 CC of SERMs such as tissue and receptor specificity. The method can also be
 CC applied to the fractionation of mixtures of SERMs to determine which
 CC components are producing agonistic and antagonistic activity. The method
 CC may be used with other receptors (e.g. progesterone, androgen,
 CC glucocorticoid, thyroid, vitamin D, beta-adrenergic, dopamine and
 CC epidermal growth factor, to identify, characterise and classify
 CC modulators of receptor activity. Peptides comprising a LXXLL motif may
 CC be used to modulate the oestrogen receptor in treating e.g. breast and
 CC ovarian cancer and ameliorating the effects of menopause, including
 CC osteoporosis. AAY65439 to AAY65652 represent oestrogen receptor,
 CC estradiol receptor and oestrogen response element binding peptides given
 CC in the exemplification of the present invention. AAZ35740 to AAZ35745
 CC represent oligonucleotides used in the exemplification of the present
 CC invention.
 CC
 XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 18; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 Db 9 GLSI 12

Search completed: August 20, 2003, 12:33:42
 Job time : 9.79518 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 12:40:51 ; Search time 4.33735 Seconds
(without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-29
Perfect score: 18
Sequence: 1 GLSI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_MA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*\n2: /cgn2_6/ptodata/1/pubpaa/PCIT_NEW_PUB.pep:*\n3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*\n4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*\n5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*\n6: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep:*\n7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*\n8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*\n9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*\n10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*\n11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*\n12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*\n13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*\n14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*\n15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*\n16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*\n17: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*\n18: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	4	11	US-09-300-425B-29
2	18	100.0	10	11	US-09-572-404B-2168
3	18	100.0	12	12	US-10-190-082-599
4	18	100.0	15	10	US-09-953-510-48
5	18	100.0	15	10	US-09-953-510-48
6	18	100.0	15	12	US-10-147-255-48
7	18	100.0	15	12	US-10-147-255-48
8	18	100.0	16	10	US-09-922-261-99
9	18	100.0	16	10	US-09-922-261-99
10	18	100.0	16	11	US-09-813-333-66
11	18	100.0	16	11	US-09-880-748-2174
12	18	100.0	16	11	US-09-880-748-2186
13	18	100.0	16	11	US-09-880-748-2186
14	18	100.0	16	12	US-09-991-225-61
15	18	100.0	16	12	US-09-991-225-61
16	18	100.0	16	12	US-09-991-225-62
17	18	100.0	16	12	US-09-991-225-62
18	18	100.0	16	12	US-09-991-225-73

16	18	100.0	16	12	US-09-991-225-74	Sequence 74, Appl
17	18	100.0	16	14	US-10-044-703-66	Sequence 66, Appl
18	18	100.0	17	15	US-10-225-567A-2048	Sequence 2048, Ap
19	18	100.0	20	11	US-09-814-604-26	Sequence 26, Appl
20	18	100.0	20	11	US-09-922-226-12	Sequence 12, Appl
21	18	100.0	20	11	US-09-922-226-165	Sequence 165, App
22	18	100.0	25	9	US-09-864-761-41951	Sequence 41951, A
23	18	100.0	25	12	US-10-036-988-3	Sequence 3, Appl1
24	18	100.0	25	12	US-09-991-225-12	Sequence 12, Appl
25	18	100.0	25	12	US-09-991-225-31	Sequence 31, Appl
26	18	100.0	27	9	US-09-864-761-48997	Sequence 48997, A
27	18	100.0	28	9	US-09-864-761-48997	Sequence 48997, A
28	18	100.0	29	9	US-09-864-761-43280	Sequence 43280, A
29	18	100.0	30	9	US-09-864-761-43280	Sequence 43280, A
30	18	100.0	32	15	US-10-106-698-7868	Sequence 7868, Ap
31	18	100.0	33	9	US-09-281-717-34	Sequence 34, Appl
32	18	100.0	33	9	US-09-281-717-34	Sequence 34, Appl
33	18	100.0	39	15	US-09-764-860-550	Sequence 550, App
34	18	100.0	42	8	US-08-424-550B-537	Sequence 537, App
35	18	100.0	42	8	US-09-864-761-48544	Sequence 48544, A
36	18	100.0	43	11	US-09-895-298-119	Sequence 119, App
37	18	100.0	44	9	US-09-764-853-439	Sequence 439, App
38	18	100.0	46	10	US-09-764-877-1059	Sequence 1059, Ap
39	18	100.0	50	10	US-09-764-847-726	Sequence 726, App
40	18	100.0	50	10	US-09-764-877-1063	Sequence 1063, Ap
41	18	100.0	50	11	US-09-809-391-575	Sequence 575, App
42	18	100.0	50	15	US-10-092-154-726	Sequence 726, App
43	18	100.0	51	11	US-09-809-391-421	Sequence 421, App
44	18	100.0	52	9	US-09-925-299-892	Sequence 892, App
45	18	100.0	52	10	US-09-922-261-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1
US-09-300-425B-29
Sequence 29, Application US/09300425B
Publication No. US20030045681A1
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARLI, Lorenzo
APPLICANT: VITI, Francesca
APPLICANT: BICCHER, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
FILE REFERENCE: SCH-1733P1
CURRENT APPLICATION NUMBER: US/09/300,425B
CURRENT FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 09/075,338
PRIOR FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
OTHER INFORMATION: antibody clone
US-09-300-425B-29
Query Match 100.0%; Score 18; DB 11; Length 4;
Best local Similarity 100.0%; Pred. No. 4.4e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2

US-09-572-404B-2168
; Sequence 2168, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 2168
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in PMEL17 OR D12S53E at 164-173 and may interact
US-09-572-404B-2168

Query Match
Best Local Similarity 100.0%; Score 18; DB 11; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
||||
Db 6 GLSI 9

RESULT 3
US-10-190-082-599
; Sequence 599, Application US/10190082
; Publication No. US20030148264A1
; GENERAL INFORMATION:
; APPLICANT: lasky, Lawrence A.
; APPLICANT: Sidhu, Sandeep S.
; APPLICANT: Held, Heike A.
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS
; FILE REFERENCE: P1905R1
; CURRENT APPLICATION NUMBER: US/10/190,082
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,634
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 683
; SEQ ID NO 599
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-190-082-599

Query Match
Best Local Similarity 100.0%; Score 18; DB 12; Length 12;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
||||
Db 3 GLSI 6

RESULT 4
US-09-953-510-48
; Sequence 48, Application US/09953510
; Patent No. US20020131975A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Marcus A.
; TITLE OF INVENTION: Abundant Extracellular
; Products and Methods for Their Production and Use
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Kurt A. Maclean
; STREET: 2029 Century Park East, Suite 3800

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/953,510
FILING DATE: 14-Sep-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-953-510-48
Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 15;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
||||
Db 10 GLSI 13

RESULT 5
US-09-953-510-49
; Sequence 49, Application US/09953510
; Patent No. US20020131975A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Marcus A.
; TITLE OF INVENTION: Abundant Extracellular
; Products and Methods for Their Production and Use
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Kurt A. Maclean
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/953,510
FILING DATE: 14-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-953-510-49
Query Match 100.0%; Score 18; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSI 4
DB 5 GSI 8
RESULT 6
US-10-147-255-48
Sequence 48, Application US/10147255
Publication No. US20030152584A1
GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular
Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/147,255
FILING DATE: 15-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/226,539A

FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-147-255-48
Query Match 100.0%; Score 18; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSI 4
DB 10 GSI 13
RESULT 7
US-10-147-255-49
Sequence 49, Application US/10147255
Publication No. US20030152584A1
GENERAL INFORMATION:
APPLICANT: Horwitz, Marcus A.
TITLE OF INVENTION: Abundant Extracellular
Products and Methods for Their Production and Use
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kurt A. Maclean
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/147,255
FILING DATE: 15-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/226,539A
FILING DATE: 15-May-2002
CLASSIFICATION: <Unknown>
APPLICATION NUMBER: US 08/447,398
FILING DATE: 23-MAY-1995
APPLICATION NUMBER: US 08/289,667
FILING DATE: 12-AUG-1994
APPLICATION NUMBER: US 08/156,358

FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maclean, Kurt A.
REGISTRATION NUMBER: 31,118
REFERENCE/DOCKET NUMBER: 112-272
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
STRAIN: Erdman
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-10-147-255-49

Query Match 100.0%; Score 18; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLSI 4
Db 5 GLSI 8

RESULT 8
US-09-922-261-99
Sequence 99, Application US/09922261
Patent No. US2002011471A1
GENERAL INFORMATION:
APPLICANT: COGENET NEUROSCIENCE, INC.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kastur C.
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
TITLE OF INVENTION: CELL DEATH
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/922,261
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US/09/461,697
PRIOR FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 99
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-261-99

Query Match 100.0%; Score 18; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 3 GLSI 6

RESULT 9
US-09-813-333-66
Sequence 66, Application US/09813333

Patent No. US20020119160A1
GENERAL INFORMATION:
APPLICANT: Degroot, Anne S
TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
FILE REFERENCE: 17999-004 US
CURRENT APPLICATION NUMBER: US/09/813,333
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,834
PRIOR FILING DATE: 2000-03-20
NUMBER OF SEQ ID NOS: 81
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 66
LENGTH: 16
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-813-333-66

Query Match 100.0%; Score 18; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 6 GLSI 9

RESULT 10
US-09-880-748-2174
Sequence 2174, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2174
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2174

Query Match 100.0%; Score 18; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 3 GLSI 6

RESULT 11
US-09-880-748-2186
Sequence 2186, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PF523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2186
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2186

Query Match 100.0%; Score 18; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
||||
DB 3 GLSI 6

RESULT 12
US-09-880-748-2980
Sequence 2980, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: P523
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2980
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2980

Query Match 100.0%; Score 18; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
||||
DB 3 GLSI 6

RESULT 13
US-09-991-225-61
Sequence 61, Application US/09991225
Publication No. US20030153063A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY11, EXPRESSED HI
TITLE OF INVENTION: HEART AND VARIANTS THEREOF
FILE REFERENCE: D0075.NP

CURRENT APPLICATION NUMBER: US/09/991,225
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 60/249,613
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/257,611
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/305,818
PRIOR FILING DATE: 2001-07-16
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Patentln version 3.0
SEQ ID NO 61
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-991-225-61

Query Match 100.0%; Score 18; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
||||
DB 11 GLSI 14

RESULT 14
US-09-991-225-62
Sequence 62, Application US/09991225
Publication No. US20030153063A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY11, EXPRESSED
FILE REFERENCE: D0075.NP
CURRENT APPLICATION NUMBER: US/09/991,225
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 60/249,613
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/257,611
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/305,818
PRIOR FILING DATE: 2001-07-16
NUMBER OF SEQ ID NOS: 81
SOFTWARE: Patentln version 3.0
SEQ ID NO 62
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-991-225-62

Query Match 100.0%; Score 18; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
||||
DB 8 GLSI 11

RESULT 15
US-09-991-225-73
Sequence 73, Application US/09991225
Publication No. US20030153063A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY11, EXPRESSED
FILE REFERENCE: D0075.NP
CURRENT APPLICATION NUMBER: US/09/991,225
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 60/249,613
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/257,611

; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-991-225-73

Query Match 100.0%; Score 18; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 GLSI 4
1111
Db 11 GLSI 14

Search completed: August 20, 2003, 13:16:45
Job time : 4.33735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 1.15663 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-29

Perfect score: 18

Sequence: 1 GLSI 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/paa/PCIT_NEW.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	100.0	6	1	PCT-US02-27760-730 Sequence 730, App
2	18	100.0	9	1	PCT-US02-27760-142 Sequence 142, App
3	18	100.0	9	1	PCT-US02-27760-350 Sequence 350, App
4	18	100.0	9	1	PCT-US02-27760-406 Sequence 406, App
5	18	100.0	9	1	PCT-US02-27760-524 Sequence 524, App
6	18	100.0	10	1	PCT-US02-27760-157 Sequence 157, App
7	18	100.0	10	1	PCT-US02-27760-196 Sequence 196, App
8	18	100.0	10	1	PCT-US02-27760-275 Sequence 275, App
9	18	100.0	10	1	PCT-US02-27760-377 Sequence 377, App
10	18	100.0	10	1	PCT-US02-27760-394 Sequence 394, App
11	18	100.0	10	1	PCT-US02-27760-456 Sequence 456, App
12	18	100.0	10	6	US-10-462-850-2168 Sequence 2168, App
13	18	100.0	23	6	US-10-411-206A-11 Sequence 11, App1
14	18	100.0	26	6	US-10-631-402-2489 Sequence 2489, App
15	18	100.0	26	6	US-10-631-441-2489 Sequence 2489, App
16	18	100.0	46	5	US-09-516-061C-43 Sequence 43, App1
17	18	100.0	48	6	US-10-612-783-6551 Sequence 6551, App
18	18	100.0	48	6	US-10-286-887-2654 Sequence 2654, App
19	18	100.0	52	6	US-10-258-898A-2654 Sequence 2654, App
20	18	100.0	57	7	US-60-485-450-1132 Sequence 1132, App
21	18	100.0	63	6	US-10-372-003A-81 Sequence 81, App1
22	18	100.0	70	6	US-10-603-113-25906 Sequence 25906, A
23	18	100.0	71	6	US-10-603-113-23802 Sequence 23802, A
24	18	100.0	73	6	US-10-603-113-22401 Sequence 22401, A
25	18	100.0	76	6	US-10-603-113-6279 Sequence 6279, App
26	18	100.0	78	6	US-10-603-113-27030 Sequence 27030, App

27	18	100.0	79	6	US-10-293-244-3272 Sequence 3272, App
28	18	100.0	79	6	US-10-603-113-26666 Sequence 26666, A
29	18	100.0	81	6	US-10-603-113-21166 Sequence 21166, A
30	18	100.0	81	6	US-10-603-113-23903 Sequence 23903, A
31	18	100.0	83	6	US-10-603-113-21820 Sequence 21820, A
32	18	100.0	85	6	US-10-603-113-23917 Sequence 23917, A
33	18	100.0	87	6	US-10-603-113-5074 Sequence 5074, App
34	18	100.0	91	6	US-10-273-573-5684 Sequence 5684, App
35	18	100.0	92	6	US-10-603-113-22075 Sequence 22075, A
36	18	100.0	93	6	US-10-630-590-2 Sequence 2, App1
37	18	100.0	93	6	US-10-630-590-68 Sequence 68, App1
38	18	100.0	94	6	US-10-372-003A-79 Sequence 79, App1
39	18	100.0	95	6	US-10-603-113-17866 Sequence 17866, A
40	18	100.0	95	6	US-10-630-590-141 Sequence 141, App
41	18	100.0	96	6	US-10-603-113-21294 Sequence 21294, A
42	18	100.0	96	6	US-10-603-113-25585 Sequence 25585, A
43	18	100.0	98	6	US-10-630-590-142 Sequence 142, App
44	18	100.0	99	6	US-10-630-590-202 Sequence 202, App
45	18	100.0	100	6	US-10-603-113-14336 Sequence 14336, A

ALIGNMENTS

RESULT 1
PCT-US02-27760-730
Sequence 730, Application PC/TUS0227760
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Chailita-Eld, Pia M.
APPLICANT: Raitano, Arthur B.
APPLICANT: Faris, Mary
APPLICANT: Hubert, Rene S.
APPLICANT: Jakobovits, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
FILE REFERENCE: 511582004340
CURRENT APPLICATION NUMBER: PCT/US02/27760
CURRENT FILING DATE: 2003-07-21
PRIOR APPLICATION NUMBER: US 60/316,664
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 736
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 730
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-27760-730
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 6;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GLSI 4
Db 1 GLSI 4
RESULT 2
PCT-US02-27760-142
Sequence 142, Application PC/TUS0227760
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Chailita-Eld, Pia M.
APPLICANT: Raitano, Arthur B.
APPLICANT: Faris, Mary
APPLICANT: Hubert, Rene S.
APPLICANT: Jakobovits, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
FILE REFERENCE: 511582004340
CURRENT APPLICATION NUMBER: PCT/US02/27760
CURRENT FILING DATE: 2003-07-21

;; PRIOR APPLICATION NUMBER: US 60/316,664
;; PRIOR FILING DATE: 2001-08-31
;; NUMBER OF SEQ ID NOS: 736
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 142
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US02-27760-142

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
|||
Db 6 GLSI 9

RESULT 3
PCT-US02-27760-350
; Sequence 350, Application PC/TUS0227760
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Chailita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 511582004340
; CURRENT APPLICATION NUMBER: PCT/US02/27760
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US 60/316,664
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-27760-350

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
|||
Db 5 GLSI 8

RESULT 4
PCT-US02-27760-406
; Sequence 406, Application PC/TUS0227760
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Chailita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 511582004340
; CURRENT APPLICATION NUMBER: PCT/US02/27760
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US 60/316,664
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406

;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US02-27760-406

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
|||
Db 3 GLSI 6

RESULT 5
PCT-US02-27760-524
; Sequence 524, Application PC/TUS0227760
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Chailita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 511582004340
; CURRENT APPLICATION NUMBER: PCT/US02/27760
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US 60/316,664
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 524
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-27760-524

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
|||
Db 5 GLSI 8

RESULT 6
PCT-US02-27760-157
; Sequence 157, Application PC/TUS0227760
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Chailita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 511582004340
; CURRENT APPLICATION NUMBER: PCT/US02/27760
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US 60/316,664
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-27760-157

Query Match 100.0%; Score 18; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 1 GLSI 4

RESULT 7
 PCT-US02-27760-196

; Sequence 196, Application PC/TUS0227760
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Challita-Eld, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; FILE REFERENCE: 511582004340
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: PCT/US02/27760
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 736
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 196
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-27760-196

Query Match 100.0%; Score 18; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 7 GLSI 10

RESULT 8
 PCT-US02-27760-275

; Sequence 275, Application PC/TUS0227760
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Challita-Eld, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; FILE REFERENCE: 511582004340
 ; CURRENT APPLICATION NUMBER: PCT/US02/27760
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: US 60/316,664
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 736
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 275
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-27760-275

Query Match 100.0%; Score 18; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4

DB ||||
 1 GLSI 4

RESULT 9
 PCT-US02-27760-377

; Sequence 377, Application PC/TUS0227760
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Challita-Eld, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; FILE REFERENCE: 511582004340
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: PCT/US02/27760
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 736
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 377
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-27760-377

Query Match 100.0%; Score 18; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 7 GLSI 10

RESULT 10
 PCT-US02-27760-394

; Sequence 394, Application PC/TUS0227760
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Challita-Eld, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; FILE REFERENCE: 511582004340
 ; CURRENT APPLICATION NUMBER: PCT/US02/27760
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: US 60/316,664
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 736
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 394
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-27760-394

Query Match 100.0%; Score 18; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 1 GLSI 4

RESULT 11

PCT-US02-27760-456
; Sequence 456, Application PC/TUS0227760
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 511582004340
; CURRENT APPLICATION NUMBER: PCT/US02/27760
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US 60/316,664
; NUMBER OF SEQ ID NOS: 2001-08-31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 456
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-27760-456

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GLSI 4
|||||
DB 3 GLSI 6

RESULT 12
US-10-462-850-2168
; Sequence 2168, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 2168
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in PMEL17 OR D12S53E at 164-173 and may interact
; OTHER INFORMATION: Sequence 2167 in this patent.
US-10-462-850-2168

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
|||||
DB 6 GLSI 9

RESULT 13
US-10-411-206A-11
; Sequence 11, Application US/10411206A
; GENERAL INFORMATION:
; APPLICANT: Irving W. WAINER et al.
; TITLE OF INVENTION: COMPUTER-BASED MODEL FOR IDENTIFICATION AND CHARACTERIZATION FOR
; TITLE OF INVENTION: NON-COMPETITIVE INHIBITORS OF NICOTINIC ACETYLCHOLINE RECEPTORS
; FILE REFERENCE: 1173-1025P
; CURRENT APPLICATION NUMBER: US/10/411,206A
; CURRENT FILING DATE: 2003-11-04

NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Table 1 Beta 1 Sequence - Transmembrane domain of ligand gated
US-10-411-206A-11

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
|||||
DB 4 GLSI 7

RESULT 14
US-10-631-402-2489
; Sequence 2489, Application US/10631402
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,402
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 2489
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Prostate
; NAME/KEY: SIGNAL
; LOCATION: -21...-1
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.50
; OTHER INFORMATION: seq SILHCSVCLFLC/OY
US-10-631-402-2489

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
|||||
DB 7 GLSI 10


```
RESULT 15
US-10-631-441-2489
; Sequence 2489, Application US/10631441
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Ducloer, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,441
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 2489
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Prostate
; NAME/KEY: SIGNAL
; LOCATION: -21..-1
; OTHER INFORMATION: Von Heljne matrix
; OTHER INFORMATION: score 4.50
; OTHER INFORMATION: seq SIFHCVCVCLFLC/QY
US-10-631-441-2489

Query Match          100.0%; Score 18; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. NO. 59;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GLSI 4
        ||||
DB       7 GLSI 10
```

Search completed: August 20, 2003, 12:45:13
Job time : 2.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using SW model

Run on: August 20, 2003, 12:25:56 ; Search time 2.79518 Seconds
(without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-29

Perfect score: 18

Sequence: 1 GLST 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR76:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	17	2 S57519	T cell receptor be
2	18	100.0	31	2 G82839	hypothetical prote
3	18	100.0	34	2 T19865	hypothetical prote
4	18	100.0	34	2 T23074	hypothetical prote
5	18	100.0	34	2 T26210	hypothetical prote
6	18	100.0	34	2 T27560	hypothetical prote
7	18	100.0	34	2 S68648	major glycoprotein
8	18	100.0	35	2 T11812	hypothetical prote
9	18	100.0	46	2 S11913	probable nitrogen
10	18	100.0	50	2 S78726	protein YKL162c-a
11	18	100.0	53	2 S56727	calcium-dependent
12	18	100.0	55	2 AB2181	hypothetical prote
13	18	100.0	56	2 C90120	40S ribosomal prot
14	18	100.0	56	2 H95182	hypothetical prote
15	18	100.0	56	2 C98050	hypothetical prote
16	18	100.0	59	2 A43601	lysine-RNA ligase
17	18	100.0	61	2 S25779	hypothetical prote
18	18	100.0	62	2 AF3295	hypothetical prote
19	18	100.0	65	2 AF1082	hypothetical prote
20	18	100.0	65	2 B89948	hypothetical prote
21	18	100.0	66	2 T03015	hypothetical prote
22	18	100.0	67	2 T44093	hypothetical prote
23	18	100.0	67	2 A12190	hypothetical prote
24	18	100.0	69	2 G95378	hypothetical prote
25	18	100.0	69	2 AF1064	probable regulator
26	18	100.0	70	2 S09281	fixu protein - Rhl
27	18	100.0	70	2 S32893	pefi protein - Sal
28	18	100.0	71	2 T01720	3a protein - canin
29	18	100.0	71	2 S01739	hypothetical prote

30	18	100.0	72	2 A60007	hypothetical prote
31	18	100.0	73	2 A42508	F14L protein - vac
32	18	100.0	73	2 T30790	hypothetical prote
33	18	100.0	73	2 AD0925	hypothetical prote
34	18	100.0	75	2 B90530	conserved hypotet
35	18	100.0	78	2 B44479	collagen alpha 1(X
36	18	100.0	79	2 T04176	ribosomal protein
37	18	100.0	79	2 AH0619	probable damage-in
38	18	100.0	83	2 C86629	hypothetical prote
39	18	100.0	83	2 C31684	hypothetical prote
40	18	100.0	85	2 A10316	sec-independent pr
41	18	100.0	86	2 H49846	spas protein - Shl
42	18	100.0	88	1 CCBM6	cytochrome c6 - gr
43	18	100.0	90	2 T04575	hypothetical prote
44	18	100.0	92	2 T12859	hypothetical prote
45	18	100.0	92	2 E96034	hypothetical prote

ALIGNMENTS

RESULT 1

S57519
T cell receptor beta chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57519
R:Burosov, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.
Submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diverse
A:Reference number: S57494
A:Accession: S57519
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-17 <BUR>
A:Cross-references: EMBL:Z49330; NID:987494; PIDN:CAA90176.1; PID:987495
C:Keywords: T-cell receptor

Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 17;
Matches 4; Conservative 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLST 4
DB 6 GLST 9

RESULT 2

G82839
hypothetical protein XF0163 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82839
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-31 <STM>
A:Cross-references: GB:AE003870; GB:AE003849; NID:9104955; PIDN:AAF62976.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.
Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Pereira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, M.C.; Franco, M.C.; Fr
J.O.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A.;Nuthors, da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva
M.; Tshahao, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0163

Query Match 100.0%; Score 18; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLSI 4
Db 10 GLSI 13

RESULT 3
T19865
hypothetical protein C40H5.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T19865
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19189
A:Accession: T19865
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <WIL>
A:Cross-references: EMBL:Z81482; PIDN:CAB03952.1; GSPDB:GN00028; CESP:C40H5.1
A:Experimental source: clone C40H5
C:Genetics:
A:Gene: CESP:C40H5.1
A:Map position: X

Query Match 100.0%; Score 18; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 23 GLSI 26

RESULT 4
T23074
hypothetical protein H12D21.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23074; T27557
R:McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19671
A:Accession: T23074
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <WIL>
A:Cross-references: EMBL:Z92849; PIDN:CAB07428.1; GSPDB:GN00023; CESP:H12D21.1
A:Experimental source: clone H12D21
R:Ainscough, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z20387
A:Accession: T27557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <WIL>
A:Cross-references: EMBL:Z78067; PIDN:CAB01526.1; GSPDB:GN00023; CESP:ZC412.7
A:Experimental source: clone ZC412
C:Genetics:
A:Gene: CESP:H12D21.1; CESP:ZC412.7
A:Map position: 5

Query Match 100.0%; Score 18; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLSI 4
Db 23 GLSI 26

RESULT 5
T26210
hypothetical protein W06A7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26210
R:Ainscough, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z20173
A:Accession: T26210
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <WIL>
A:Cross-references: EMBL:Z78066; PIDN:CAB01519.1; GSPDB:GN00023; CESP:W06A7.5
A:Experimental source: clone W06A7
C:Genetics:
A:Gene: CESP:W06A7.5
A:Map position: 5

Query Match 100.0%; Score 18; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 23 GLSI 26

RESULT 6
T27560
hypothetical protein ZC412.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27560
R:Ainscough, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z20387
A:Accession: T27560
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <WIL>
A:Cross-references: EMBL:Z78067; PIDN:CAB01529.1; GSPDB:GN00023; CESP:ZC412.6
A:Experimental source: clone ZC412
C:Genetics:
A:Gene: CESP:ZC412.6
A:Map position: 5

Query Match 100.0%; Score 18; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLSI 4
Db 23 GLSI 26

RESULT 7
S68648
major glycoprotein PSP-I - pig (fragments)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 17-Mar-1999
C:Accession: S68648
R:Calvete, J.J.; Postolova, Z.; Sanz, L.; Adermann, K.; Thole, H.H.; Toepfer-Petersen
FEBS Lett. 379, 207-211, 1996

A:Title: Mapping the heparin-binding domain of boar spermadhesins.
A:Reference number: S68648; MUID:96184566; PMID:8603690
A:Accession: S68648
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7;8-12;13-20;21-27;28-32;33-34 <CAL>

Query Match 100.0%; Score 18; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 23 GLSI 26

RESULT 8

hypothetical protein ORF35 - Norway spruce chloroplast
C:Species: chloroplast Picea abies (Norway spruce)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: T11812
R:Klumper, S.; Kanka, S.; Rlesner, D.; Etscheid, M.
submitted to the EMBL Data Library, March 1997
A:Description: Characterisation of a Norway spruce chloroplast DNA clone: Complete nucle
A:Reference number: Z17349
A:Accession: T11812
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-35 <KLU>
A:Cross-references: EMBL:U92462; NID:92959581; PID:92959588
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 100.0%; Score 18; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 18 GLSI 21

RESULT 9

probable nitrogen regulation protein II (EC 2.7.3.-) - Xanthomonas campestris pv. campe
C:Species: Xanthomonas campestris pv. campestris
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 17-Jul-1998
C:Accession: S11913
R:Osborn, A.E.; Clarke, B.R.; Stevens, B.J.H.; Daniels, M.J.
Mol. Gen. Genet. 222, 145-151, 1990
A:Title: Use of oligonucleotide probes to identify members of two-component regulatory s
A:Reference number: S11912; MUID:91042416; PMID:2233675
A:Accession: S11913
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-46 <OSB>
C:Superfamily: glnL regulatory protein II: sensor histidine kinase homology
C:Keywords: ATP; autophosphorylation; phosphohistidine; phosphoprotein; phosphotransfer
F:1-41/Domain: sensor histidine kinase homology (fragment) <SHK>

Query Match 100.0%; Score 18; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 9 GLSI 12

RESULT 10
S78726

protein YKL162c-a - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae
C:Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 15-Jan-1999
C:Accession: S78726
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hllger, F.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S77976
A:Accession: S78726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <VA2>
A:Cross-references: EMBL:Z28163; MIPS:YKL162c-a
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 11L

Query Match 100.0%; Score 18; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 14 GLSI 17

RESULT 11

556727
calcium-dependent protein kinase (EC 2.7.1.-) (clone AspK7) - oat (fragment)
N:Alternate names: calmodulin-independent protein kinase
C:Species: Avena sativa (oat)
C:Date: 27-Oct-1995 #sequence_revision 19-Jan-1996 #text_change 29-Sep-1999
C:Accession: S56727
R:Hutley, A.K.; Phillips, A.L.
Plant Mol. Biol. 27, 1043-1052, 1995
A:Title: Gibberellin-regulated expression in oat aleurone cells of two kinases that s
A:Reference number: S56638; MUID:95284341; PMID:7766874
A:Accession: S56727
A:Molecule type: mRNA
A:Residues: 1-53 <HUT>
C:Superfamily: calcium-dependent protein kinase: calmodulin repeat homology; protein
C:Keywords: ATP; EF hand; phosphotransferase; protein kinase
F:1-53/Domain: protein kinase homology (fragment) <KIN>

Query Match 100.0%; Score 18; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 19 GLSI 22

RESULT 12

AB2181
hypothetical protein asr3001 [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AB2181
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001

Query Match 100.0%; Score 18; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 19 GLSI 22

RESULT 12
AB2181
hypothetical protein asr3001 [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AB2181
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2181
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAW4700.1; PID:gt132095; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asr3001

Query Match 100.0%; Score 18; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GLSI 4
 ||||
 Db 36 GLSI 39

RESULT 13

C90120
 40S ribosomal protein S29A [imported] - Guillardia theta nucleomorph
 C:Species: nucleomorph Guillardia theta
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Aug-2002
 C:Accession: C90120
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reif
 Nature 410, 1091-1096, 2001
 A:Title: The highly reduced genome of an enslaved algal nucleus.
 A:Reference number: A99082; PMID:11323671; PMID:11323671
 A:Accession: C90120
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-56 <DOU>
 A:Cross-references: GB:AF083031; NID:913794279; PIDN:AAK39656.1; GSPDB:GN00152
 C:Genetics:
 A:Gene: rps29A
 A:Map position: 3
 A:Genome: nucleomorph
 C:Superfamily: Escherichia coli ribosomal protein S14
 C:Keywords: nucleomorph

Query Match 100.0%; Score 18; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GLSI 4
 ||||
 Db 35 GLSI 38

RESULT 14

H95182
 Hypothetical protein SP1570 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: H95182
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtapple,
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smilh, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; PMID:21357209; PMID:11463916
 A:Accession: H95182
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-56 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75657.1; PID:914973061; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPI570

Query Match 100.0%; Score 18; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GLSI 4
 ||||
 Db 31 GLSI 34

RESULT 15

C98050
 Hypothetical protein spr1428 [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: C98050
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zock, C.; Baltz, R.H.; Jaskunas, S
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; PMID:21429245; PMID:11544234
 A:Accession: C98050
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-56 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAL00232.1; PID:915459082; GSPDB:GN00174
 C:Genetics:
 A:Gene: spr1428

Query Match 100.0%; Score 18; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GLSI 4
 ||||
 Db 31 GLSI 34

Search completed: August 20, 2003, 12:42:18
 Job time : 5.79518 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 1.51807 Seconds
(without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-29

Perfect score: 18
Sequence: 1 GLST 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	70	1	FIXU_RHILF
2	18	100.0	71	1	FIXU_RHILF
3	18	100.0	73	1	FIXU_RHILF
4	18	100.0	73	1	FIXU_RHILF
5	18	100.0	75	1	FIXU_RHILF
6	18	100.0	80	1	FIXU_RHILF
7	18	100.0	83	1	FIXU_RHILF
8	18	100.0	85	1	FIXU_RHILF
9	18	100.0	86	1	FIXU_RHILF
10	18	100.0	88	1	FIXU_RHILF
11	18	100.0	94	1	FIXU_RHILF
12	18	100.0	96	1	FIXU_RHILF
13	18	100.0	96	1	FIXU_RHILF
14	18	100.0	96	1	FIXU_RHILF
15	18	100.0	96	1	FIXU_RHILF
16	18	100.0	97	1	FIXU_RHILF
17	18	100.0	99	1	FIXU_RHILF
18	18	100.0	99	1	FIXU_RHILF
19	18	100.0	100	1	FIXU_RHILF
20	18	100.0	101	1	FIXU_RHILF
21	18	100.0	101	1	FIXU_RHILF
22	18	100.0	101	1	FIXU_RHILF
23	18	100.0	101	1	FIXU_RHILF
24	18	100.0	106	1	FIXU_RHILF
25	18	100.0	108	1	FIXU_RHILF
26	18	100.0	108	1	FIXU_RHILF
27	18	100.0	112	1	FIXU_RHILF
28	18	100.0	112	1	FIXU_RHILF
29	18	100.0	113	1	FIXU_RHILF
30	18	100.0	113	1	FIXU_RHILF
31	18	100.0	114	1	FIXU_RHILF
32	18	100.0	114	1	FIXU_RHILF
33	18	100.0	115	1	FIXU_RHILF

34	18	100.0	118	1	MOBC_THIFE
35	18	100.0	118	1	YK26_RALST
36	18	100.0	118	1	YK26_RALST
37	18	100.0	119	1	YK26_RALST
38	18	100.0	121	1	YK26_RALST
39	18	100.0	122	1	YK26_RALST
40	18	100.0	124	1	YK26_RALST
41	18	100.0	128	1	YK26_RALST
42	18	100.0	131	1	YK26_RALST
43	18	100.0	133	1	YK26_RALST
44	18	100.0	133	1	YK26_RALST
45	18	100.0	133	1	YK26_RALST

ALIGNMENTS

RESULT 1
FIXU_RHILF
ID: P42710.1995 (Rel. 32, Created)
AC: 01-NOV-1995 (Rel. 32, Last sequence update)
DT: 01-NOV-1995 (Rel. 32, Last sequence update)
DE: 16-OCT-2001 (Rel. 40, Last annotation update)
Protein fixu.
GN: Rhizobium leguminosarum (biovar trifolii).
OS: Rhizobium leguminosarum (biovar trifolii).
OC: Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC: Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX: NCBI_TaxID=386;
RN: [1]
RP: Rhizobium leguminosarum (biovar trifolii).
RC: MEDLINE=90158123; PubMed=2622339;
RX: Iismaa S.E., Ealing P.M., Scott K.F., Watson J.M.;
RA: "Molecular linkage of the nif/fix and nod gene regions in Rhizobium
RT: leguminosarum biovar trifolii.";
RL: Mol. Microbiol. 3:1753-1764(1989).
CC: -1- SIMILARITY: TO K.PNEUMONIAE AND A.VINELANDII NIFT PROTEINS AND TO
THE N-TERMINAL OF R.MELIOTI MOSH PROTEIN.
CC: This SWISS-PROT entry is copyright. It is produced through a collaboration
CC: between the Swiss Institute of Bioinformatics and the EMBL outstation
CC: the European Bioinformatics Institute. There are no restrictions on its
CC: use by non-profit institutions as long as its content is in no way
CC: modified and this statement is not removed. Usage by and for commercial
CC: entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC: or send an email to license@sib-sib.ch).
DR: EMBL; X51963; CAB37405.1; -
DR: PIR; S09281; S09281.
KW: Nitrogen fixation; Plasmid.
SQ: SEQUENCE 70 AA; 7810 MM; FG45370F284FACBE CRC64;
Query Match
Best local Similarity 100.0%; Score 18; DB 1; Length 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GLST 4
DB 11 GLST 14
RESULT 2
FIXU_RHILF
ID: P42710.1995 (Rel. 32, Created)
AC: 01-NOV-1988 (Rel. 09, Created)
DT: 01-NOV-1988 (Rel. 09, Last sequence update)
DE: 15-SEP-2003 (Rel. 42, Last annotation update)
DE: X2A protein.
OS: Porcine transmissible gastroenteritis coronavirus (strain Purdue)

OS (TGEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11151;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88078100; PubMed=2825819;
 RA Rasmussen D., Gelfi J., Laude H.;
 RT "Enteric coronavirus TGEV: partial sequence of the genomic RNA, its
 organization and expression."
 RL Biochimie 69:591-600(1987).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X06371; CAA29670.1; -
 DR PIR: S01739; S01739.
 DR InterPro: IPR006784; Corona_3.
 DR Pfam: PF04694; Corona_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 71 AA; 7710 MW; 433A511590A3A83A CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 71;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLST 4
 DB 58 GLST 61
 RESULT 3
 VF14_VACCC STANDARD; PRT; 73 AA.
 ID VF14_VACCC
 AC P21019;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein F14.
 GN F14L.
 OS Vaccinia virus (strain Copenhagen).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10249;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021027; PubMed=2219722;
 RA Geibel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RT "The complete DNA sequence of vaccinia virus."
 RL Virology 179:247-266(1990).
 RN [2]
 RP COMPLETE GENOME.
 RA Geibel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RT "Appendix to 'The complete DNA sequence of vaccinia virus'."
 RL Virology 179:517-563(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M35027; AAA48032.1; -
 DR PIR: A42508; A42508.

SQ SEQUENCE 73 AA; 8293 MW; 0C8422D6BEC0E27C CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 73;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLST 4
 DB 9 GLST 12
 RESULT 4
 VF14_VACCP STANDARD; PRT; 73 AA.
 ID VF14_VACCP
 AC P29890;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Protein F14.
 GN F14L OR F4.
 OS Vaccinia virus (strain L-1VP).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=31531;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mikryukov N.N., Chizhikov V.E., Prikhod'ko G.G., Urmmanov I.M.,
 RA Serpinskii O.I., Blinov V.M., Nikulin A.E., Vasilenko S.K.;
 RT "Structural-functional organization of segment of vaccinia virus
 genome."
 RL Biotechnologia 4:442-449(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M57977; AAA48283.1; -
 DR PIR: M57977; M57977.
 SQ SEQUENCE 73 AA; 8307 MW; 089422D6B9C7F27B CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 73;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLST 4
 DB 9 GLST 12
 RESULT 5
 Y146_MYCPU STANDARD; PRT; 75 AA.
 ID Y146_MYCPU
 AC Q98R64;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein MYPU_1460.
 GN MYPU_1460.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=UAB CT1P;
 MEDLINE=21267165; PubMed=11353084;
 RA Chaudhary I., Helling R., Ferris S., Barbe V., Samson D., Gallisson F.,
 RA Moszer I., Dpyvlg K., Wroblewski H., Viari A., Rocha E.P.C.,
 RA Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen
 Mycoplasma pulmonis.";

RL Nucleic Acids Res. 29:2145-2153(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPE0154 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL445563; CAC13319.1; -
 DR PIR: B90530; B90530.
 DR MyPULst; MYPU_1460; -
 DR HAMAP: MF_00363; -; 1.
 DR InterPro: IPR005359; UPE0154.
 DR Pfam: PF03672; UPE0154; 1.
 DR ProDom: PD048972; UPE0154; 1.
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 8 28 POTENTIAL.
 SQ SEQUENCE 75 AA; 8470 MW; A92264859CE7E908 CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 75;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 DB 12 GLSI 15
 RESULT 6
 LIPE_RAT ID LIPE_RAT STANDARD; PRT; 80 AA.
 AC Q8VBX1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Endothelial lipase (EC 3.1.1.3) (Endothelial-derived lipase)
 DE (EDL) (Fragment).
 GN LIPE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BN-Lx/Cub; and SHR/Olaipcv;
 RX MEDLINE=21913086; PubMed=11924532;
 RA Boone A.C.M., den Bieman M.G., Van Lith H., van Zutphen B.F.M.;
 RT "Sequencing and chromosomal assignment of the rat endothelial-derived
 RT lipase gene (Lipg)."
 RL DNA Seq. 12:285-287(2001).
 CC -1- FUNCTION: Has phospholipase and triglyceride lipase activities.
 CC Hydrolyses high density lipoproteins (HDL) more efficiently than
 CC other lipoproteins. Binds heparin (By similarity).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
 CC fatty acid anion.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- MISCELLANEOUS: It is termed endothelial lipase due to the fact
 CC that it is synthesized in endothelial cells, a characteristic that
 CC distinguishes it from other members of the family. However this
 CC protein is also expressed in other cell types.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AY027561; AAK14774.1; -
 DR EMBL: AY027562; AAK14775.1; -
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00151; Lipase; 1.
 KM Hydrolase; Lipid degradation; Heparin-binding.
 FT NON_TER 1 1
 FT NON_TER 80 80
 SQ SEQUENCE 80 AA; 8791 MW; 460120E3CDD29F5D CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 DB 16 GLSI 19
 RESULT 7
 Y102_CVBM ID Y102_CVBM STANDARD; PRT; 83 AA.
 AC P22054;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein IORF2.
 OS Bovine coronavirus (strain Mebus) (BCOV) (BCV), and
 OS Bovine coronavirus (strain Quebec) (BCOV) (BCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11132, 11133;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mebus;
 RX MEDLINE=90204700; PubMed=2319653;
 RA Klenzle T.E., Abraham S., Hogue B.G., Brian D.A.;
 RT "Structure and orientation of expressed bovine coronavirus
 RT hemagglutinin-esterase protein."
 RL J. Virol. 64:1834-1838(1990).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mebus;
 RX MEDLINE=91353426; PubMed=2103108;
 RA Klenzle T.E., Abraham S., Hogue B.G., Brian D.A.;
 RT "Structure and expression of the bovine coronavirus hemagglutinin
 RT protein."
 RL Adv. Exp. Med. Biol. 276:95-102(1990).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Quebec;
 RX MEDLINE=89279187; PubMed=2732684;
 RA Parker M.D., Cox G.J., Dereg D., Fitzpatrick D.R., Babluk L.A.;
 RT "Cloning and in vitro expression of the gene for the E3
 RT haemagglutinin glycoprotein of bovine coronavirus."
 RL J. Gen. Virol. 70:155-164(1989).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M30612; AAA66395.1; -
 DR EMBL: S50936; AAB19564.1; -
 DR PIR: C31684; C31684.
 KM Hypothetical protein.
 SQ SEQUENCE 83 AA; 9512 MW; 7343P2AC9330DAE9 CRC64;
 Query Match 100.0%; Score 18; DB 1; Length 83;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
1111
Db 43 GLSI 46

RESULT 8

TATE_YERPE STANDARD; PRT; 85 AA.

ID TATE_YERPE

AC 082DHL;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Sec-independent protein translocase (update)

GN TATE OR IP02597 OR Y1170.

OS Yersinia pestis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OX NCBI_TaxID=632;

[1]

SEQUENCE FROM N.A.

RC STRAIN=CO-92 / Biovar Orientalis;

RA MEDLINE=21470413; PubMed=11586360;

RA Parthill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,

RA Baker S., Basham D., Bentley S.D., Brooks R., Cerdano-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,

RA Leachner S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,

RA "Genome sequence of Yersinia pestis, the causative agent of plague,"

RA Nature 413:523-527(2001).

[2]

SEQUENCE FROM N.A.

RC STRAIN=KIM5 / Biovar Mediaevalis;

RA MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,

RA Perry R.D., "Genome sequence of Yersinia pestis KIM,"

RA J. Bacteriol. 184:4601-4611(2002).

[1]

FUNCTION: Required for correct localization of precursor proteins

bearing signal peptides with the twin arginine conserved motif

S/T-R-R-F-L-R. This sec-independent pathway is termed TAT for

twin-arginine translocation system. This system mainly transports

proteins with bound cofactors that require folding prior to export

(by similarity).

-1- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).

-1- SIMILARITY: BELONGS TO THE TATA/E FAMILY.

[1]

SEQUENCE FROM N.A.

RC STRAIN=PROT / Biovar Prototypus;

RA MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,

RA Perry R.D., "Genome sequence of Yersinia pestis PROT,"

RA J. Bacteriol. 184:4601-4611(2002).

[1]

FUNCTION: Required for correct localization of precursor proteins

bearing signal peptides with the twin arginine conserved motif

S/T-R-R-F-L-R. This sec-independent pathway is termed TAT for

twin-arginine translocation system. This system mainly transports

proteins with bound cofactors that require folding prior to export

(by similarity).

-1- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).

-1- SIMILARITY: BELONGS TO THE TATA/E FAMILY.

[1]

SEQUENCE FROM N.A.

RC STRAIN=PROT / Biovar Prototypus;

RA MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,

RA Perry R.D., "Genome sequence of Yersinia pestis PROT,"

RA J. Bacteriol. 184:4601-4611(2002).

[1]

FUNCTION: Required for correct localization of precursor proteins

bearing signal peptides with the twin arginine conserved motif

S/T-R-R-F-L-R. This sec-independent pathway is termed TAT for

twin-arginine translocation system. This system mainly transports

proteins with bound cofactors that require folding prior to export

(by similarity).

-1- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).

-1- SIMILARITY: BELONGS TO THE TATA/E FAMILY.

[1]

SEQUENCE FROM N.A.

RC STRAIN=PROT / Biovar Prototypus;

RA MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,

RA Perry R.D., "Genome sequence of Yersinia pestis PROT,"

RA J. Bacteriol. 184:4601-4611(2002).

[1]

FUNCTION: Required for correct localization of precursor proteins

bearing signal peptides with the twin arginine conserved motif

S/T-R-R-F-L-R. This sec-independent pathway is termed TAT for

twin-arginine translocation system. This system mainly transports

proteins with bound cofactors that require folding prior to export

(by similarity).

-1- SUBCELLULAR LOCATION: Inner-membrane bound (Probable).

Query Match 100.0%; Score 18; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
1111
Db 3 GLSI 6

RESULT 9

SPAO_SHIFL

ID SPAO_SHIFL

AC P40705; 055297;

DT 01-FEB-1995 (Rel. 31, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Surface presentation of antigens protein spa9 (Spa9 protein).

GN SPAO OR SPA9 OR CP0154.

OS Shigella flexneri, and

OS Shigella sonnei.

OC Plasmid pMR100, plasmid pMYSH6000, and plasmid pCP301.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Shigella.

OX NCBI_TaxID=623, 624;

[1]

SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=PMR100;

RA MEDLINE=20566792; PubMed=11115111;

RA Buchrieser C., Glaeser P., Rusnock C., Nedjari H., d'Hauteville H.,

RA Kunst F., Sansonetti P., Parsot C.,

RA "The virulence plasmid pMR100 and the repertoire of proteins secreted

by the type III secretion apparatus of Shigella flexneri,"

RA Mol. Microbiol. 38:760-771(2000).

[2]

SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=PMR100;

RA MEDLINE=21189246; PubMed=11292750;

RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grothbeck E.J., Burland V.,

RA Blattner F.R.,

RA "Complete DNA sequence and analysis of the large virulence plasmid of

Shigella flexneri,"

RA Infect. Immun. 69:3271-3285(2001).

[3]

SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=YSH6000 / Serotype 2a; PLASMID=PMYSH6000;

RA MEDLINE=93224456; PubMed=8385666;

RA Sasakawa C., Komatsu K., Tope T., Suzuki T., Yoshikawa M.,

RA "Eight genes in region 5 that form an operon are essential for

invasion of epithelial cells by Shigella flexneri 2a,"

RA J. Bacteriol. 175:2334-2346(1993).

[4]

SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a; PLASMID=PCP301;

RA MEDLINE=22272406; PubMed=12384590;

RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,

RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,

RA Yu J.,

RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity

through comparison with genomes of Escherichia coli K12 and O157,"

RA Nucleic Acids Res. 30:4432-4441(2002).

[5]

SEQUENCE OF 1-53 FROM N.A.

RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=PMR100;

RA MEDLINE=92193289; PubMed=1312536;

RA Venkatesan M.M., Buysse J.M., Oaks E.V.,

RA "Surface presentation of Shigella flexneri invasion plasmid antigens

requires the products of the spa locus,"

RA J. Bacteriol. 174:1990-2001(1992).

[6]

SEQUENCE FROM N.A.

RC SPECIES=S.sonnei; STRAIN=HM383;

RA Atakawa, Kato J. I., Ito K. I., Marubane H.:
RT "Comparison and high conservation of nucleotide sequences of spa-mxi
RT regions between S.somnei and S.flexneri -- Identification of a new
RT gene coding plausible membrane protein.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: REQUIRED FOR SURFACE PRESENTATION OF INVASION PLASMID
CC ANTIGENS. COULD PLAY A ROLE IN PRESERVING THE TRANSLOCATION
CC COMPETENCE OF THE IPA ANTIGENS. REQUIRED FOR INVASION AND FOR
CC SECRETION OF THE THREE IPA PROTEINS.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC -1- SIMILARITY: BELONGS TO THE FliG/MOPD/SPA FAMILY.
CC
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC

CC
DR EMBL; AL391753; CAC05829.1; -
DR EMBL; AF348706; AAK18473.1; -
DR EMBL; D13663; BAA02830.1; -
DR EMBL; AF386526; AAL72556.1; -
DR EMBL; M81458; -; NOT_ANNOTATED_CDS.
DR EMBL; D50601; BAA09163.1; ALT_INIT.
DR PIR; H49846; H49846.
DR InterPro: IPR002191; Bac_export_3.
DR InterPro: IPR006306; HrpO.
DR Pfam; PF03313; Bac_export_3; 1.
DR PRINTS; PRO0952; TYPE3IMOPROT.
DR TIGRFAMs; TIGR01403; fliQ_rel_III; 1.
KW Plasmid; Virulence; Transmembrane.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 54 74 POTENTIAL.
SQ SEQUENCE 86 AA; 9429 MW; 552D3DBC9C896E3 CRC64;

Query Match	100.0%	Score 18:	DB 1:	length 86;
Best Local Similarity	100.0%	Pred. NC:	2.8e+02;	
Matches 4; Conservative	0;	Mismatches	0;	Gaps 0

Qy	1	GLST	4
Db	30	GLST	33

CC6_BRYMA			RESULT 10
ID_CYC6_BRYMA	STANDARD:	PRT:	88 AA.
AC	P11448;		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-		
DE	553).		
GN	PETY.		
OS	Bryopsis maxima (Green alga).		
OC	Eukaryota: Viridiplantae: Chlorophyta; Ulvophyceae; Caulerpaceles;		
OC	Bryopsidaceae: Bryopsids.		
OX	NCBI_TaxID=3129;		
RN	[1]		
RP	SEQUENCE, AND CHARACTERIZATION.		
RX	MEDLINE=88139277; PubMed3481367;		
RA	*Okamoto Y., Minami Y., Matsubara H., Sugimura Y.;		
RT	"Studies on algal cytochromes VI: some properties and amino acid		
RT	sequence of cytochrome c6 from a green alga, Bryopsis maxima.";		
RL	J. Biochem. 102:1251-1260(1987).		
CC	-1- FUNCTION: Functions as an electron carrier between membrane-bound		
CC	cytochrome b6f and photosystem I in oxygenic photosynthesis.		
CC	-1- SUBUNIT: Monomer.		
CC	-1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.		
CC	-1- PTM: Blinds one heme group per molecule.		
CC	-1- MISCELLANEOUS: The midpoint redox potential for this protein is		

CC	+385 MV.
CC	-I- SMILIARITY: Belongs to the cytochrome c family. Pcty subfamily.
DR	PIR: A30021; CCBM6.
DR	HSSP: P08197; ICRY.
DR	InterPro: IPRO003088; Cyt_C1.
DR	InterPro: IPRO02329; Cyt_C1C.
DR	InterPro: IPRO02323; Cyt_C1C.
DR	InterPro: IPRO00345; CytC_heme_bind.
DR	Pfam: PF00034; Cytochrome_c_1.
DR	PRINTS: PR00605; CYTOCHROME.CIC.
DR	PROSITE: PR00607; CYTOCHROME.CIC.
DR	PROSITE: PS00190; CYTOCHROME_C_1.
KW	Electron transport; Photosynthesis; Heme; Chloroplast; Thylakoid.
FT	BINDING 15 HEME (COVALENT).
FT	BINDING 18 HEME (COVALENT).
FT	METAL 19 IRON (HEME AXIAL LIGAND).
FT	METAL 61 IRON (HEME AXIAL LIGAND).
SQ	SEQUENCE 88 AA: 9286 MW: EFDECC2580175E01 CRC64;

Query Match	100.0%	Score 18;	DB 1;	Length 88;
Best Local Similarity	100.0%	Pred. No. 2.3e+02;		
Matches	4;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	1	GLSI	4
Db	44	GLSI	47

ID	SPDA_STRLI	STANDARD:	PRT:	94 AA.
AC	P22407.			
DT	01-AUG-1991	(Rel. 19, Created)		
DT	01-AUG-1991	(Rel. 19, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Protein spda.			
GN	SPDA.			
OS	Streptomyces lividans.			
OG	Plasmid pIJ101.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_Taxid=1916;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89008081; PubMed=3170481;			
RA	Kendall K.J., Cohen S.N.;			
RT	Complete nucleotide sequence of the Streptomyces lividans plasmid			
RT	pIJ101 and correlation of the sequence with genetic properties."			
RL	J. Bacteriol. 170:4634-4651(1988).			
-----	-1- FUNCTION: INVOLVED IN PLASMID TRANSFER.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce ,			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL: M21778; AAA88409.1; -			
KW	PIR: F31844; F31844.			
FT	Plasmid; Transmembrane.			
FT	TRANSMEM 41 68			
SQ	SEQUENCE 94 AA; 9525 MW; 47D80CB57D824656 CRC64;	POTENTIAL.		
QY	1 GLSI 4	100.0%; Score 18; DB 1; Length 94;		
DB	83 GLSI 86	Best Local Similarity 100.0%; Pred. No. 3.1e+02;		
		Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0		

CC	enables i-egouts to	-----	
CC	or send an email to	license@sb.slb.ch).	
CC		-----	
CC	EMBL, M21778; AA868409.1; -		
DR	PIR; F31844; F31844.		
KW	plasmid; Transmembrane.		
FT	TRANSMEM	41	POTENTIAL.
FT	SEQUENCE	94 AA; 9535 MW;	47D80CB57D824656 CRC64;

Query Match	100.0%	Score 18;	DB 1;	Length 94;
Best Local Similarity	100.0%	Pred No.3	le+02;	
Matches 4;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	GLSI	4
	1	1	1
Db	83	GLSI	86

```

RESULT 12
GRP7 DAUCA STANDARD: PRT: 96 AA.
ID GRP7 DAUCA
AC P37704;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE Glycine-rich protein DC7.1 precursor.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RA Aleith F., Richter G.;
RT "Gene expression during induction of somatic embryogenesis in carrot
RT cell suspensions.";
RT Planta 183:17-24(1990).
CC -I- FUNCTION: MAY BE CONNECTED WITH THE INITIATION OF EMBRYOGENESIS
CC OR WITH THE METABOLIC CHANGES PRODUCED BY THE REMOVAL OF AUXINS.
CC -I- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED DURING EARLY
CC EMBRYOGENESIS.
CC -I- INDUCTION: BY THE REMOVAL OF AUXINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X15706; CAA3736.1; -
DR PIR: S35715; S35715.
KW Repeat: signal.
FT SIGNAL 1 25
FT CHAIN 26 96 POTENTIAL.
FT DOMAIN 42 67 GLYCINE-RICH PROTEIN DC7.1.
FT REPEAT 42 50 2 APPROXIMATE REPEATS OF H-H-G(4,6)-H.
FT REPEAT 61 67 1.
SQ SEQUENCE 96 AA: 9319 MW: 7C00DD4637B7A364 CRC64;
Query Match 100.0%; Score 18; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 10 GLSI 13

RESULT 13
NUML CERCA STANDARD: PRT: 96 AA.
ID NUML CERCA
AC Q34049;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN ND4L.
OS Ceratitis capitata (Mediterranean fruit fly).
OC Ceratitridion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Tephritidae; Tephritidae; Ceratitis.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Guatemala Laboratory colony;
RX MEDLINE=95261546; PubMed=7742977;
RA Gasparich G.E., Shepard W.S., Han H.Y., McPherson B.A., Steck G.J.;

```

```

RT RT
RT RT "Analysis of mitochondrial DNA and development of PCR-based
RT diagnostic molecular markers for Mediterranean fruit fly (Ceratitid
RT capitata) populations.";
RT Insect Mol. Biol. 4:61-67(1995).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U12925; AAA85798.1; -
DR InterPro: IPR001133; Oxidored_4L.
DR InterPro: IPR003214; Oxidred4L.
DR Pfam: PF00420; Oxidored_Q2; 1.
DR ProDom: PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 96 AA: 11346 MW: B743137D69FP2331 CRC64;
Query Match 100.0%; Score 18; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
Db 74 GLSI 77

RESULT 14
NUML DROME STANDARD: PRT: 96 AA.
ID NUML DROME
AC P18934;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MT:ND4L OR ND4L.
OS Drosophila melanogaster (Fruit fly).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Phytroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bretagne;
RX MEDLINE=88212147; PubMed=3130291;
RA Garesse R.;
RT "Drosophila melanogaster mitochondrial DNA: gene organization and
RT evolutionary considerations.";
RT Genetics 118:649-663(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R, and Zimbabwe 53;
RX MEDLINE=20363871; PubMed=10903372;
RA Ballard J.W.O.;
RT "Comparative genomics of mitochondrial DNA in members of the
RT Drosophila melanogaster subgroup.";
RT J. Mol. Evol. 51:48-63(2000).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M37275; AAA69712.1; -

```

DR EMBL: U37541; AAC47820.1; -
 DR EMBL: AF200828; AAF77235.1; -
 DR EMBL: AF200829; AAF77247.1; -
 DR PIR: S01188; S01188.
 DR FlyBase: FBgn0013683; mt:ND4L.
 DR InterPro: IPR001133; Oxidored_4L.
 DR InterPro: IPR003214; Oxidred4L.
 DR Pfam: PF00420; Oxidored_g2; 1.
 DR ProDom: PD000359; Oxidred4L; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 96 AA: 11359 MW: D38357D738A175B CRC64;

Query_Match 100.0%; Score 18; DB 1; Length 96;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLSI 4
 ||||
 DB 74 GLSI 77

RESULT 15

NUM_DROVA

ID NUM_DROVA STANDARD: PRT: 96 AA.

AC P07708.1

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).

GN ND4L.

OS Drosophila yakuba (Fruit fly).

OC Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7245;

ON [1]

RP SEQUENCE FROM N.A.

RC STRAIN=2317.6 Ivory Coast;

RX MEDLINE=86089137; PubMed=3001325;

RA Clary D.O., Wolstenholme D.R.;

RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide

sequence, gene organization, and genetic code.";

RL J. Mol. Evol. 22:252-271(1985).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=84221393; PubMed=6328435;

RA Clary D.O., Wahlthner J.A., Wolstenholme D.R.;

RT "Sequence and arrangement of the genes for cytochrome b, URF1, URF4L,

URF4, URF5, URF6 and five tRNAs in Drosophila mitochondrial DNA.";

RL Nucleic Acids Res. 12:3747-3762(1984).

CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL: X03240; CAA26994.1; -

CC PIR: A30020; A30020.

CC FlyBase: FBgn0013186; Dyak\mt:ND4L.

CC InterPro: IPR001133; Oxidored_4L.

CC InterPro: IPR003214; Oxidred4L.

CC Pfam: PF00420; Oxidored_g2; 1.

CC ProDom: PD000359; Oxidred4L; 1.

CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

KW SEQUENCE 96 AA: 11386 MW: 59ABEC7D738A174B CRC64;

SQ

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GLSI 4
 ||||
 DB 74 GLSI 77
 Search completed: August 20, 2003, 12:34:53
 Job time : 2.62918 secs

Query Match 100.0%; Score 18; DB 1; Length 96;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;

THIS PAGE BLANK (USPTO)

XX Claim 39; Page 195; 202pp; English.
XX
PS
CC AAY15618-Y15759 represent peptides used to make the fluorogenic or
CC fluorescent reporter molecules of the invention. These molecules
CC contain a peptide moiety (e.g. present sequence) which acts as a
CC substrate for enzymes involved in apoptosis or protease or peptidase
CC enzymes. The compounds can be used as fluorogenic or fluorescent
CC substrates for enzymes. Depending on the peptide moiety used, the
CC fluorescent molecules can be used for detecting or measuring the
CC activity of an enzyme involved in the apoptosis cascade in cells; to
CC determine whether a test compound has an effect on an enzyme involved
CC in the apoptosis cascade in cells; for determining the sensitivity of
CC an animal with cancer to treatment with chemotherapeutic agents or
CC determining whether a test substance inhibits, prevents, causes or
CC enhances cell death of test cells; for detecting or measuring the
CC activity of a viral protease in cells; for determining whether a test
CC compound has an effect on the activity of a viral protease in cells;
CC and for measuring the activity or determining whether a test substance
CC has an effect on the activity of a protease or peptidase in cells.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 20; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFSF 4
1111
1 SFSF 4
Db
RESULT 2
AAY80902
ID AAY80902 standard; peptide; 5 AA.
XX
AC AAY80902;
XX
DT 22-MAY-2000 (first entry)
XX
DE Fluorophore-labelled protease substrate peptide, SEQ ID NO:131.
XX
KW Protease substrate; fluorescent label; fluorophore; rhodamine;
KW blocking group; halobenzoyl group; cleavage; caspase; viral protease;
KW methionine aminopeptidase type 2; MetAP-2; drug screening.
XX
OS Synthetic.
XX
PN WO200004914-A1.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US16423.
XX
PR 21-JUL-1998; 98US-0093642.
XX
PA (CYTO-) CYTOVIA INC.
PA (ZHAN/) ZHANG H.
PA (CAIS/) CAI S X.
PA (DREW/) DREWE J A.
PA (YANG/) YANG W.
XX
PI Zhang H, Cai SX, Drewe JA, Yang W;
XX
DR WPI; 2000-195079/17.
XX
PT New fluorescently labeled amino acids or peptides, used as substrates
PT for detecting enzymes or their modulators, e.g. anticancer or antiviral
PT agents, contains a halobenzoyl N-blocking group
XX
XX Claim 29; Page 109; 174pp; English.
PS The invention relates to fluorescently labelled peptides containing
CC

CC a halobenzoyl group on the fluorophore. They are of the structure
CC peptide-Y-Z, where Z represents a halo-substituted benzoyl blocking
CC group, Y is a fluorescent or fluorogenic moiety (preferably a
CC rhodamine), and the peptide-Y bond is cleavable by the enzyme being
CC assayed. The labelled peptides are reporters for detecting intracellular
CC proteolytic enzymes, particularly caspases and other enzymes involved in
CC apoptosis; viral proteases (e.g., HIV, herpes simplex virus-1, human
CC cytomegalovirus and hepatitis C virus proteases); and methionine
CC aminopeptidase type 2 (MetAP-2). The peptides are particularly used to
CC identify modulators of these enzymes which may be potentially useful as
CC agents for treating conditions such as cancer, neurodegeneration,
CC autoimmune diseases, myocardial infarction and viral infection.
CC Modulators identified may also be used to prolong the life of cells being
CC cultured for recombinant protein production or to monitor the treatment
CC of cancer with chemotherapeutic agents. Inhibitors of MetAP-2 are
CC potential anti- angiogenic or anticancer agents. Sequences
CC AAY80782-Y80910 represent peptides, some of which are specifically
CC claimed, which may be used in assay methods according to the invention.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 20; DB 21; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFSF 4
1111
1 SFSF 4
Db
RESULT 3
AAY15739
ID AAY15739 standard; Peptide; 6 AA.
XX
AC AAY15739;
XX
DT 27-JUL-1999 (first entry)
XX
DE Peptide used to make fluorescent reporter molecules.
XX
KW Fluorogenic; fluorescent reporter molecule; enzyme substrate;
KW apoptosis; protease; peptidase; apoptosis cascade; cancer;
KW chemotherapeutic agent; cell death; viral protease activity.
XX
OS Synthetic.
XX
PN WO9918856-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21231.
XX
PF 03-MAR-1998; 98US-0033661.
XX
PR 10-OCT-1997; 97US-0061582.
XX
PA (CYTO-) CYTOVIA INC.
XX
PI Cai SX, Drewe JA, Keana JFW, Weber E, Zhang H;
XX
DR WPI; 1999-312448/26.
XX
PT New fluorogenic or fluorescent reporter molecules
XX
PS Claim 39; Page 193; 202pp; English.
XX
CC AAY15618-Y15759 represent peptides used to make the fluorogenic or
CC fluorescent reporter molecules of the invention. These molecules
CC contain a peptide moiety (e.g. present sequence) which acts as a
CC substrate for enzymes involved in apoptosis or protease or peptidase
CC enzymes. The compounds can be used as fluorogenic or fluorescent
CC substrates for enzymes. Depending on the peptide moiety used, the
CC fluorescent molecules can be used for detecting or measuring the
CC activity of an enzyme involved in the apoptosis cascade in cells; to

CC determine whether a test compound has an effect on an enzyme involved
 CC in the apoptosis cascade in cells; for determining the sensitivity of
 CC an animal with cancer to treatment with chemotherapeutic agents or
 CC determining whether a test substance inhibits, prevents, causes or
 CC enhances cell death of test cells; for detecting or measuring the
 CC activity of a viral protease in cells; for determining whether a test
 CC compound has an effect on the activity of a viral protease in cells;
 CC and for measuring the activity or determining whether a test substance
 CC has an effect on the activity of a protease or peptidase in cells.

XX Sequence 6 AA:

Query Match 100.0%; Score 20; DB 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 1 SFSF 4

RESULT 4

ID AAY80893 standard; peptide; 6 AA.

AC AAY80893;

DT 22-MAY-2000 (first entry)

DE Fluorophore-labelled protease substrate peptide, SEQ ID NO:122.

XX Protease substrate; fluoroscent label; fluorophore; rhodamine;
 KW blocking group; halobenzoyle group; cleavage; caspase; viral protease;
 KM methionine aminopeptidase type 2; MetAP-2; drug screening.

OS Synthetic.

PN WO200004914-A1.

PD 03-FEB-2000.

PF 21-JUL-1999; 99WO-US16423.

PR 21-JUL-1998; 98US-0093642.

PA (CYTO-) CYTOVIA INC.

PA (ZHAN/) ZHANG H.

PA (CAIS/) CAI S X.

PA (DREW/) DREWE J A.

PA (YANG/) YANG W.

PI Zhang H, Cai SX, Drewe JA, Yang W;

DR WPI; 2000-195079/17.

XX New fluorescently labeled amino acids or peptides, used as substrates
 PT for detecting enzymes or their modulators, e.g. anticancer or antiviral
 PT agents, contains a halobenzoyle N-blocking group -

XX Clalm 29; Page 109; 174pp; English.

XX The invention relates to fluorescently labelled peptides containing
 CC a halobenzoyle group on the fluorophore. They are of the structure
 CC peptide-Y-2, where 2 represents a halo-substituted benzoyl blocking
 CC group, Y is a fluorescent or fluorogenic moiety (preferably a
 CC rhodamine), and the peptide-Y bond is cleavable by the enzyme being
 CC assayed. The labelled peptides are reporters for detecting intracellular
 CC proteolytic enzymes, particularly caspases and other enzymes involved in
 CC apoptosis; viral proteases (e.g., HIV, herpes simplex virus-1, human
 CC cytomegalovirus and hepatitis C virus proteases); and methionine
 CC aminopeptidase type 2 (MetAP-2). The peptides are particularly used to
 CC identify modulators of these enzymes which may be potentially useful as
 CC agents for treating conditions such as cancer, neurodegeneration,

CC autoimmune diseases, myocardial infarction and viral infection.
 CC Modulators identified may also be used to prolong the life of cells being
 CC cultured for recombinant protein production, or to monitor the treatment
 CC of cancer with chemotherapeutic agents. Inhibitors of MetAP-2 are
 CC potential anti- angiogenic or anticancer agents. Sequences
 CC AAY80782-Y80910 represent peptides, some of which are specifically
 CC claimed, which may be used in assay methods according to the invention.

XX Sequence 6 AA:

Query Match 100.0%; Score 20; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 1 SFSF 4

RESULT 5

ID AAO17262 standard; peptide; 6 AA.

AC AAO17262;

DT 08-JUL-2002 (first entry)

DE A thaliana receptor kinase RKS4 peptide fragment.

XX Vegetative propagation; plant; phytochrome; receptor kinase-like SERK;
 KW RKS.

OS Arabidopsis thaliana.

PN EP1094113-A1.

PD 25-APR-2001.

PF 22-OCT-1999; 99EP-0203480.

PR 22-OCT-1999; 99EP-0203480.

PA (GENE-) GENETWISTER TECHNOLOGIES BV.

PI Schmidt EDL, Van Der Kop DAM, De Boer AD;

DR WPI; 2002-228902/29.

XX In vitro culture propagation of a plant from plant starting material,
 PT comprises stimulating root/shoot initiation by introducing a
 PT recombinant gene product into the starting material, thus reducing
 PT phytochrome addition to culture -

XX Disclosure; Page 72; 171pp; English.

XX The present invention relates to a culture method for propagating a plant
 CC from a plant starting material, where root or shoot initiation is
 CC stimulated by introducing a gene into the starting material which allows
 CC the reduction or absence of phytochrome addition to the culture. The
 CC method is used for the propagation (preferably, seedless propagation) of
 CC a plant from a plant starting material in an in vitro culture method.
 CC Nucleic acids encoding receptor-like kinases are useful in the method.
 CC The present sequence is a fragment of a receptor-like kinase protein.

XX Sequence 6 AA:

Query Match 100.0%; Score 20; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 2 SFSF 5

DT	10-JUN-2003	(first entry)
DE	Staphylococcus aureus	CHIPS-related peptide #103.
XX	CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5ar;	
KW	formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;	
KW	inflammation; cardiovascular disease; central nervous system disease;	
KW	gastrointestinal disease; skin disease; genitourinary disease;	
KW	joint disease; respiratory disease; HIV infection; antiinflammatory;	
KW	cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;	
KW	gynecological; immunosuppressive; anti-HIV.	
XX	Staphylococcus aureus.	
OS	Synthetic.	
XX	WO2003006048-A1.	
PN	23-JAN-2003.	
XX	11-JUL-2001; 2001WO-EP08004.	
XX	11-JUL-2001; 2001WO-EP08004.	
PF	11-JUL-2001; 2001WO-EP08004.	
XX	11-JUL-2001; 2001WO-EP08004.	
PR	11-JUL-2001; 2001WO-EP08004.	
XX	(JARI-) JARI PHARM BV.	
PA	Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtz JAW;	
XX	Van Strijp JAG;	
PI	WPI; 2003-247783/25.	
DR	Combination of peptides derived from chemotaxis inhibiting protein from	
XX	Staphylococcus aureus (CHIPS) having CHIPS activity, useful in	
PT	prophylaxis and treatment of inflammation, cardiovascular, skin and	
PT	kidney diseases	
XX	-	
PS	Disclosure; Page 11; 89pp; English.	
XX	The present invention relates to peptides (ABR44811-ABR47162 and	
CC	ABR47164-ABR47185) derived from the Chemotaxis Inhibitory Protein (CHIPS)	
CC	from Staphylococcus aureus. The peptide fragments are useful in the	
CC	prophylaxis or treatment of diseases or disorders involving the	
CC	C5a-receptor (C5ar) and/or formylated peptide receptor (FPR) or	
CC	neutrophils, monocytes and endothelial cells or involving acute or	
CC	chronic inflammatory reactions. The diseases or disorders include	
CC	cardiovascular diseases, disease of the central nervous system,	
CC	gastrointestinal diseases, skin diseases, genitourinary diseases, joint	
CC	diseases, respiratory diseases and HIV infection.	
CC		
XX	Sequence 6 AA:	
SQ		
Query Match	100.0%;	Score 20; DB 24; Length 6;
Best Local Similarity	100.0%;	Pred. NO. 9.3e+05;
Matches	4;	Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 SFSE 4	
Db	2 SFSE 5	
RESULT 8		
ABR45301		
ID	ABR45301 standard; Peptide; 6 AA.	
AC	ABR45301;	
XX	10-JUN-2003 (first entry)	
DE	Staphylococcus aureus	CHIPS-related peptide #491.
XX	CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5ar;	
KW	formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;	
KW	inflammation; cardiovascular disease; central nervous system disease;	
KW	gastrointestinal disease; skin disease; genitourinary disease;	
KW	joint disease; respiratory disease; HIV infection; antiinflammatory;	
KW	cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;	
KW	gynecological; immunosuppressive; anti-HIV.	
XX	Staphylococcus aureus.	
OS	Synthetic.	
XX	WO2003006048-A1.	
PN	23-JAN-2003.	
XX	11-JUL-2001; 2001WO-EP08004.	
XX	11-JUL-2001; 2001WO-EP08004.	
PF	11-JUL-2001; 2001WO-EP08004.	
XX	11-JUL-2001; 2001WO-EP08004.	
PR	11-JUL-2001; 2001WO-EP08004.	
XX	(JARI-) JARI PHARM BV.	
PA	Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtz JAW;	
XX	Van Strijp JAG;	
PI	WPI; 2003-247783/25.	
DR	Combination of peptides derived from chemotaxis inhibiting protein from	
XX	Staphylococcus aureus (CHIPS) having CHIPS activity, useful in	
PT	prophylaxis and treatment of inflammation, cardiovascular, skin and	
PT	kidney diseases	
XX	-	
PS	Disclosure; Page 11; 89pp; English.	
XX	The present invention relates to peptides (ABR44811-ABR47162 and	
CC	ABR47164-ABR47185) derived from the Chemotaxis Inhibitory Protein (CHIPS)	
CC	from Staphylococcus aureus. The peptide fragments are useful in the	
CC	prophylaxis or treatment of diseases or disorders involving the	
CC	C5a-receptor (C5ar) and/or formylated peptide receptor (FPR) or	
CC	neutrophils, monocytes and endothelial cells or involving acute or	
CC	chronic inflammatory reactions. The diseases or disorders include	
CC	cardiovascular diseases, disease of the central nervous system,	
CC	gastrointestinal diseases, skin diseases, genitourinary diseases, joint	
CC	diseases, respiratory diseases and HIV infection.	
CC		
XX	Sequence 6 AA:	
SQ		
Query Match	100.0%;	Score 20; DB 24; Length 6;
Best Local Similarity	100.0%;	Pred. NO. 9.3e+05;
Matches	4;	Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 SFSE 4	
Db	2 SFSE 5	

KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 OS Staphylococcus aureus.
 OS Synthetic.
 XX WO2003006048-A1.
 XX
 XX 23-JAN-2003.
 PD
 PF 11-JUL-2001; 2001WO-EP08004.
 XX
 XX 11-JUL-2001; 2001WO-EP08004.
 PR
 XX 11-JUL-2001; 2001WO-EP08004.
 PA (JARI-) JARI PHARM BV.
 PI Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 XX WPI: 2003-247783/25.
 DR
 XX
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases
 XX
 XX Disclosure: Page 12; 89pp; English.
 PS
 XX
 XX The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.
 CC
 XX
 XX Sequence 6 AA:
 SQ
 Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 IIII
 DB 2 SFSF 5
 DE
 XX
 XX RESULT 9
 ID ABR45305 standard; Peptide; 6 AA.
 AC ABR45305;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 XX Staphylococcus aureus CHIPS-related peptide #495.
 DE
 XX
 XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW Formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 XX Staphylococcus aureus.
 OS Synthetic.
 OS
 XX WO2003006048-A1.
 PN

XX
 XX 23-JAN-2003.
 PD
 XX
 XX 11-JUL-2001; 2001WO-EP08004.
 PF
 XX
 XX 11-JUL-2001; 2001WO-EP08004.
 PR
 XX
 XX (JARI-) JARI PHARM BV.
 PA
 PI Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 XX WPI: 2003-247783/25.
 DR
 XX
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases
 XX
 XX Disclosure: Page 12; 89pp; English.
 PS
 XX
 XX The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.
 CC
 XX
 XX Sequence 6 AA:
 SQ
 Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 IIII
 DB 2 SFSF 5
 DE
 XX
 XX RESULT 10
 ID ABR45637 standard; Peptide; 6 AA.
 AC ABR45637;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 XX Staphylococcus aureus CHIPS-related peptide #827.
 DE
 XX
 XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW Formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 XX Staphylococcus aureus.
 OS Synthetic.
 OS
 XX WO2003006048-A1.
 PN
 XX 23-JAN-2003.
 PD
 XX 11-JUL-2001; 2001WO-EP08004.
 PR
 XX 11-JUL-2001; 2001WO-EP08004.
 PA (JARI-) JARI PHARM BV.

XX Van Kessel CPM, Gosselaar-de Haas CTC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 XX WPI; 2003-247783/25.
 XX
 XX
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases
 XX
 PS Disclosure; Page 13; 89pp; English.
 XX
 XX The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.
 CC
 XX
 XX Sequence 6 AA;
 SQ
 Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SF5F 4
 ||||
 Db 2 SF5F 5
 -

RESULT 11
 ABR45641
 ID ABR45641 standard; Peptide: 6 AA.
 XX
 AC ABR45641;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Staphylococcus aureus CHIPS-related peptide #831.
 XX
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 PN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP08004.
 XX
 PR 11-JUL-2001; 2001WO-EP08004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-de Haas CTC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 XX WPI; 2003-247783/25.
 XX
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in

PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases
 XX
 XX Disclosure; Page 13; 89pp; English.
 XX
 XX The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.
 CC
 XX
 XX Sequence 6 AA;
 SQ
 Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SF5F 4
 ||||
 Db 2 SF5F 5
 -

RESULT 12
 ABR46029
 ID ABR46029 standard; Peptide: 6 AA.
 XX
 AC ABR46029;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Staphylococcus aureus CHIPS-related peptide #1219.
 XX
 KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 PN WO2003006048-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2001; 2001WO-EP08004.
 XX
 PR 11-JUL-2001; 2001WO-EP08004.
 XX
 PA (JARI-) JARI PHARM BV.
 XX
 PI Van Kessel CPM, Gosselaar-de Haas CTC, Kruijtzer JAW;
 PI Van Strijp JAG;
 XX
 XX WPI; 2003-247783/25.
 XX
 XX Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 XX kidney diseases
 XX
 PS Disclosure; Page 15; 89pp; English.
 XX
 XX The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the

CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.

XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 IIII
 DB 2 SFSF 5

RESULT 13

ABR46033
 ID ABR46033 standard; Peptide: 6 AA.

XX
 AC ABR46033;

XX
 DT 10-JUN-2003 (first entry)

XX
 DE Staphylococcus aureus CHIPS-related peptide #1223.

XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.

OS Staphylococcus aureus.
 OS Synthetic.

XX
 PN WO2003006048-A1.

XX
 PD 23-JAN-2003.

XX
 PF 11-JUL-2001; 2001WO-EP08004.

XX
 PR 11-JUL-2001; 2001WO-EP08004.

XX
 PA (JARI-) JARI PHARM BV.

XX
 PI Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtz JAW;
 PI Van Strijp JAG;

XX
 DR WPI; 2003-247783/25.

XX
 PT Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases
 XX
 PS Disclosure: Page 15; 89pp; English.

XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.

SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 IIII
 DB 2 SFSF 5

RESULT 14

ABR46421
 ID ABR46421 standard; Peptide: 6 AA.

XX
 AC ABR46421;

XX
 DT 10-JUN-2003 (first entry)

XX
 DE Staphylococcus aureus CHIPS-related peptide #1611.

XX CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5aR;
 KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
 KW inflammation; cardiovascular disease; central nervous system disease;
 KW gastrointestinal disease; skin disease; genitourinary disease;
 KW joint disease; respiratory disease; HIV infection; antiinflammatory;
 KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
 KW gynecological; immunosuppressive; anti-HIV.

OS Staphylococcus aureus.
 OS Synthetic.

XX
 PN WO2003006048-A1.

XX
 PD 23-JAN-2003.

XX
 PF 11-JUL-2001; 2001WO-EP08004.

XX
 PR 11-JUL-2001; 2001WO-EP08004.

XX
 PA (JARI-) JARI PHARM BV.

XX
 PI Van Kessel CPM, Gosselaar-de Haas CJC, Kruijtz JAW;
 PI Van Strijp JAG;

XX
 DR WPI; 2003-247783/25.

XX
 PT Combination of peptides derived from chemotaxis inhibiting protein from
 PT Staphylococcus aureus (CHIPS) having CHIPS activity, useful in
 PT prophylaxis and treatment of inflammation, cardiovascular, skin and
 PT kidney diseases
 XX
 PS Disclosure: Page 16; 89pp; English.

XX
 CC The present invention relates to peptides (ABR44811-ABR47162 and
 CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
 CC from Staphylococcus aureus. The peptide fragments are useful in the
 CC prophylaxis or treatment of diseases or disorders involving the
 CC C5a-receptor (C5aR) and/or formylated peptide receptor (FPR) or
 CC neutrophils, monocytes and endothelial cells or involving acute or
 CC chronic inflammation reactions. The diseases or disorders include
 CC cardiovascular diseases, disease of the central nervous system,
 CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
 CC diseases, respiratory diseases and HIV infection.

XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 24; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 IIII

Db 2 SFSF 5

RESULT 15

ABR46425
ID ABR46425 standard; Peptide; 6 AA.

XX ABR46425;

DT 10-JUN-2003 (first entry)

DE Staphylococcus aureus CHIPS-related peptide #1615.

XX
KW CHIPS; Chemotaxis Inhibitory Protein; C5a-receptor; C5ar;
KW formylated peptide receptor; FPR; neutrophil; monocyte; endothelial cell;
KW inflammation; cardiovascular disease; central nervous system disease;
KW gastrointestinal disease; skin disease; genitourinary disease;
KW joint disease; respiratory disease; HIV infection; antiinflammatory;
KW cardiant; cerebroprotective; neuroprotective; nootropic; dermatological;
KW gynecological; immunosuppressive; anti-HIV.

OS Staphylococcus aureus.
OS Synthetic.

XX
FN WO2003006048-A1.

XX
PD 23-JAN-2003.

XX
PF 11-JUL-2001; 2001WO-EP08004.

XX
PR 11-JUL-2001; 2001WO-EP08004.

XX
PA (JARI-) JARI PHARM BV.

XX
PI Van Kessel CPM, Gosselaar-de Haas CTC, Kruijzer JAM;
PI Van Strijp JAG;

XX
DR WPI: 2003-247783/25.

XX
PT Combination of peptides derived from chemotaxis inhibiting protein from
PT Staphylococcus aureus (CHIPS) having CHIPS activity; useful in
PT prophylaxis and treatment of inflammation, cardiovascular, skin and
PT kidney diseases

XX
PS Disclosure; Page 16; 89pp; English.

XX
CC The present invention relates to peptides (ABR44811-ABR47162 and
CC ABR47164-ABR47385) derived from the Chemotaxis Inhibitory Protein (CHIPS)
CC from Staphylococcus aureus. The peptide fragments are useful in the
CC prophylaxis or treatment of diseases or disorders involving the
CC C5a-receptor (C5ar) and/or formylated peptide receptor (FPR) or
CC neutrophils, monocytes and endothelial cells or involving acute or
CC chronic inflammation reactions. The diseases or disorders include
CC cardiovascular diseases, disease of the central nervous system,
CC gastrointestinal diseases, skin diseases, genitourinary diseases, joint
CC diseases, respiratory diseases and HIV infection.

XX
SQ Sequence 6 AA:

Query Match 100.0%; Score 20; DB 24; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
Db 2 SFSF 5

Search completed: August 20, 2003, 12:33:43
Job time : 9.79518 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 2.77108 Seconds

(Without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-30

Sequence: 1 SFSF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCtUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	5	3	US-09-357-952-131
2	20	100.0	5	4	US-09-521-650-131
3	20	100.0	5	4	US-09-168-888-131
4	20	100.0	6	3	US-09-357-952-122
5	20	100.0	6	4	US-09-521-650-122
6	20	100.0	6	4	US-09-168-888-122
7	20	100.0	7	3	US-09-357-952-33
8	20	100.0	7	4	US-09-521-650-33
9	20	100.0	7	4	US-09-168-888-33
10	20	100.0	10	3	US-09-053-941-19
11	20	100.0	10	4	US-09-817-413-19
12	20	100.0	12	3	US-08-825-369A-1
13	20	100.0	13	3	US-08-630-915A-187
14	20	100.0	13	4	US-08-630-915A-202
15	20	100.0	14	3	US-08-825-369A-2
16	20	100.0	14	3	US-09-208-966-46
17	20	100.0	15	3	US-09-230-222-6
18	20	100.0	15	4	US-09-230-225B-13
19	20	100.0	23	4	US-09-227-357-158
20	20	100.0	25	4	US-09-227-357-624
21	20	100.0	30	3	US-08-974-549A-153
22	20	100.0	31	4	US-09-205-258-882
23	20	100.0	34	3	US-09-215-221-16
24	20	100.0	35	4	US-09-149-476-619
25	20	100.0	40	4	US-09-023-905A-31
26	20	100.0	48	4	US-09-461-325-143
27	20	100.0	50	4	US-09-205-258-431

28	20	100.0	52	3	US-08-917-299-25	Sequence 25, Appl
29	20	100.0	52	3	US-09-422-662-25	Sequence 25, Appl
30	20	100.0	52	4	US-09-730-763-25	Sequence 25, Appl
31	20	100.0	52	4	US-09-429-370-25	Sequence 25, Appl
32	20	100.0	56	4	US-09-227-357-671	Sequence 671, Appl
33	20	100.0	62	4	US-09-205-258-367	Sequence 367, Appl
34	20	100.0	63	4	US-09-187-999-29	Sequence 29, Appl
35	20	100.0	64	4	US-09-328-352-5188	Sequence 5188, Ap
36	20	100.0	66	4	US-09-328-352-7435	Sequence 7435, Ap
37	20	100.0	70	4	US-09-107-532A-6168	Sequence 6168, Ap
38	20	100.0	74	3	US-08-905-223-294	Sequence 294, Ap
39	20	100.0	75	4	US-09-107-332A-5775	Sequence 5775, Ap
40	20	100.0	78	4	US-09-198-452A-1246	Sequence 1246, Ap
41	20	100.0	81	4	US-09-482-273-262	Sequence 262, Ap
42	20	100.0	96	4	US-09-107-532A-4962	Sequence 4962, Ap
43	20	100.0	104	4	US-09-230-485-9	Sequence 9, Appl
44	20	100.0	105	4	US-08-881-189B-13	Sequence 13, Appl
45	20	100.0	114	4	US-09-107-532A-4832	Sequence 4832, Ap

ALIGNMENTS

RESULT 1
US-09-357-952-131
Sequence 131, Application US/09357952
Patent No. 6248904

GENERAL INFORMATION:

APPLICANT: Zhang, Han-Zhong

APPLICANT: Cai, Sui Xiong

APPLICANT: Yang, Wu

TITLE OF INVENTION: No. 6248904el Fluorescence Screening Assays for Caspases, Peptidases, Protease

TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Protease

TITLE OF INVENTION: Other Enzymes and the Use Thereof

FILE REFERENCE: 1735.0030001

CURRENT FILING DATE: US/09/357,952

EARLIER FILING DATE: 1999-07-21

EARLIER FILING DATE: 21-JUL-1998

NUMBER OF SEQ ID NOS: 139

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 131

LENGTH: 5

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Peptide

US-09-357-952-131

Query Match 100.0%; Score 20; DB 3; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0;

Qy 1 SFSF 4

Db 1 SFSF 4

RESULT 2

US-09-521-650-131

Sequence 131, Application US/09521650

Patent No. 6335429

GENERAL INFORMATION:

APPLICANT: Weber, Eckard

APPLICANT: Cai, Sui Xiong

APPLICANT: Keana, John F.W.

APPLICANT: Drewe, John A.

APPLICANT: Zhang, Han-Zhong

TITLE OF INVENTION: No. 6335429el Fluorescent or Fluorescent Reporter Molecules an

TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence

TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the

;; TITLE OF INVENTION: Use Thereof
;; FILE REFERENCE: 1735.0290002
;; CURRENT APPLICATION NUMBER: US/09/521,650
;; CURRENT FILING DATE: 2000-03-08
;; EARLIER APPLICATION NUMBER: 09/168,888
;; EARLIER FILING DATE: 1998-10-09
;; EARLIER APPLICATION NUMBER: US 60/061,582
;; EARLIER FILING DATE: 1997-10-10
;; EARLIER APPLICATION NUMBER: US 09/033,661
;; EARLIER FILING DATE: 1998-03-03
;; NUMBER OF SEQ ID NOS: 142
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 131
;; LENGTH: 5
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;; OTHER INFORMATION: Peptide
US-09-521-650-131

Query Match 100.0%; Score 20; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
|||
Db 1 SFSF 4

RESULT 3
US-09-168-888-131
;; Sequence 131, Application US/09168888
;; Patent No. 6342611
;; GENERAL INFORMATION:
;; APPLICANT: Weber, Eckard
;; APPLICANT: Cai, Sui Xiong
;; APPLICANT: Keana, John F.W.
;; APPLICANT: Drewe, John A.
;; APPLICANT: Zhang, Han-Zhong
;; TITLE OF INVENTION: No. 6342611 Fluorogenic or Fluorescent Reporter Molecules and
;; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
;; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
;; TITLE OF INVENTION: Use Thereof
;; FILE REFERENCE: 1735.0290002
;; CURRENT APPLICATION NUMBER: US/09/168,888
;; CURRENT FILING DATE: 1998-10-09
;; EARLIER APPLICATION NUMBER: US 60/061,582
;; EARLIER FILING DATE: 1997-10-10
;; EARLIER APPLICATION NUMBER: US 09/033,661
;; EARLIER FILING DATE: 1998-03-03
;; NUMBER OF SEQ ID NOS: 142
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 131
;; LENGTH: 5
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;; OTHER INFORMATION: Peptide
US-09-168-888-131

Query Match 100.0%; Score 20; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
|||
Db 1 SFSF 4

RESULT 4
US-09-357-952-122

;; Sequence 122, Application US/09357952
;; Patent No. 6248904
;; GENERAL INFORMATION:
;; APPLICANT: Zhang, Han-Zhong
;; APPLICANT: Cai, Sui Xiong
;; APPLICANT: Drewe, John A.
;; APPLICANT: Yang, Wu
;; TITLE OF INVENTION: No. 6248904 Fluorescence Dyes and Their Applications for Who
;; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Protea
;; TITLE OF INVENTION: Other Enzymes and the Use Thereof
;; FILE REFERENCE: 1735.0030001
;; CURRENT APPLICATION NUMBER: US/09/357,952
;; CURRENT FILING DATE: 1999-07-21
;; EARLIER APPLICATION NUMBER: US 60/093,642
;; EARLIER FILING DATE: 21-JUL-1998
;; NUMBER OF SEQ ID NOS: 139
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 122
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;; OTHER INFORMATION: Peptide
US-09-357-952-122

Query Match 100.0%; Score 20; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
|||
Db 1 SFSF 4

RESULT 5
US-09-521-650-122
;; Sequence 122, Application US/09521650
;; Patent No. 6335429
;; GENERAL INFORMATION:
;; APPLICANT: Weber, Eckard
;; APPLICANT: Cai, Sui Xiong
;; APPLICANT: Keana, John F.W.
;; APPLICANT: Drewe, John A.
;; APPLICANT: Zhang, Han-Zhong
;; TITLE OF INVENTION: No. 6335429 Fluorogenic or Fluorescent Reporter Molecules an
;; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
;; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
;; TITLE OF INVENTION: Use Thereof
;; FILE REFERENCE: 1735.0290002
;; CURRENT APPLICATION NUMBER: US/09/521,650
;; CURRENT FILING DATE: 2000-03-08
;; EARLIER APPLICATION NUMBER: 09/168,888
;; EARLIER FILING DATE: 1998-10-09
;; EARLIER APPLICATION NUMBER: US 60/061,582
;; EARLIER FILING DATE: 1997-10-10
;; EARLIER APPLICATION NUMBER: US 09/033,661
;; EARLIER FILING DATE: 1998-03-03
;; NUMBER OF SEQ ID NOS: 142
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 122
;; LENGTH: 6
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;; OTHER INFORMATION: Peptide
US-09-521-650-122

Query Match 100.0%; Score 20; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
Db 1 SFSF 4

RESULT 6

US-09-168-888-122
; Sequence 122, Application US/09168888
; Patent No. 6342611
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sul Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6342611 Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/168,888
; EARLIER FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 122
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-168-888-122

Query Match 100.0%; Score 20; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
Db 1 SFSF 4

RESULT 7

US-09-357-952-33
; Sequence 33, Application US/09357952
; Patent No. 6248904
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sul Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: No. 6248904 Fluorescence Dyes and Their Applications for Whole-
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Proteases
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/09/357,952
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US 60/093,642
; EARLIER FILING DATE: 21-JUL-1998
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 33
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-357-952-33

Query Match 100.0%; Score 20; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
Db 1 SFSF 4

RESULT 8

US-09-521-650-33
; Sequence 33, Application US/09521650
; Patent No. 6335429
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sul Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6335429 Fluorogenic or Fluorescent Reporter Molecules an
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/521,650
; EARLIER FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 09/168,888
; EARLIER FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 33
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-521-650-33

Query Match 100.0%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
Db 1 SFSF 4

RESULT 9

US-09-168-888-33
; Sequence 33, Application US/09168888
; Patent No. 6342611
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sul Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6342611 Fluorogenic or Fluorescent Reporter Molecules an
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/168,888
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
US-09-168-888-33

EARLIER FILING DATE: 1998-03-03
NUMBER OF SEQ ID NOS: 142
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 33
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-168-888-33

Query Match 100.0%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 1 SFSF 4

RESULT 10
US-09-053-941-19
Sequence 19, Application US/09053941
Patent No. 6271354
GENERAL INFORMATION:
APPLICANT: SRINIVISAN, ALGARSAMY
APPLICANT: KOPROWSKI, HILARY
TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS
FILE REFERENCE: Chimeric Viral Proteins
CURRENT APPLICATION NUMBER: US/09/053,941
CURRENT FILING DATE: 1998-04-02
EARLIER APPLICATION NUMBER: 60/043,380
EARLIER FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 19
LENGTH: 10
TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
PUBLICATION INFORMATION:
JOURNAL: Meth. Enzymol.
VOLUME: 38
PAGES: 289-
DATE: 1974
US-09-053-941-19

Query Match 100.0%; Score 20; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 2 SFSF 5

RESULT 11
US-09-817-413-19
Sequence 19, Application US/09817413
Patent No. 6436648
GENERAL INFORMATION:
APPLICANT: SRINIVISAN, ALGARSAMY
APPLICANT: KOPROWSKI, HILARY
TITLE OF INVENTION: CHIMERIC VIRAL PROTEINS
FILE REFERENCE: Chimeric Viral Proteins
CURRENT APPLICATION NUMBER: US/09/817,413
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/043,380
PRIOR FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 19
LENGTH: 10

TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
PUBLICATION INFORMATION:
JOURNAL: Meth. Enzymol.
VOLUME: 38
PAGES: 289-
DATE: 1974
US-09-817-413-19

Query Match 100.0%; Score 20; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 2 SFSF 5

RESULT 12
US-08-825-369A-1
Sequence 1, Application US/08825369A
Patent No. 6084060
GENERAL INFORMATION:
APPLICANT: Moore
TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825,369A
FILING DATE: March 28, 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: O'Dea, Sean W.
REGISTRATION NUMBER: 37690
REFERENCE/DOCKET NUMBER: 381-21 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-825-369A-1

Query Match 100.0%; Score 20; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 5 SFSF 8

RESULT 13
US-08-630-915A-187
Sequence 187, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.

APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLER, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-187

Query Match 100.0%; Score 20; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 14
US-08-630-915A-202
Sequence 202, Application US/08630915A
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLER, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 202:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Other
OTHER INFORMATION: May or may not have carboxy-terminal
OTHER INFORMATION: amide and/or biotinylated N-terminal
US-08-630-915A-202

Query Match 100.0%; Score 20; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 15
US-08-825-369A-2
Sequence 2, Application US/08825369A
Patent No. 6084060
GENERAL INFORMATION:
APPLICANT: MOORE
TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
TITLE OF INVENTION: PROGENITOR CELLS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825,369A
FILING DATE: March 28, 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: O'Dea, Sean W.
REGISTRATION NUMBER: 37690
REFERENCE/DOCKET NUMBER: 381-21 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-825-369A-2

Query Match 100.0%; Score 20; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
|||
Db 5 SFSF 8

Search completed: August 20, 2003, 12:44:20
Job time : 4.77108 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51; Search time 4.33735 Seconds

(Without alignments)
121.698 Million cell updates/secTitle: US-09-512-082-30
Perfect score: 20
Sequence: 1 SFSF 4Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5Searched: 497079 seqs, 131961718 residues
Total number of hits satisfying chosen parameters: 497079Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published_Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCIT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCITUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	11	US-09-300-425B-30
2	20	100.0	5	10	US-09-947-387-131
3	20	100.0	6	10	US-09-947-387-122
4	20	100.0	7	10	US-09-947-387-33
5	20	100.0	11	9	US-09-205-658-293
6	20	100.0	12	10	US-09-934-251A-1
7	20	100.0	13	9	US-09-879-957-187
8	20	100.0	13	9	US-09-879-957-202
9	20	100.0	14	10	US-09-934-251A-2
10	20	100.0	14	11	US-09-775-052-46
11	20	100.0	15	11	US-09-892-877-413
12	20	100.0	15	11	US-09-948-783-413
13	20	100.0	22	16	US-10-260-212-3
14	20	100.0	23	11	US-09-983-802-158
15	20	100.0	23	15	US-10-097-065-345

16	20	100.0	24	9	US-09-864-761-44483	Sequence 44483, A
17	20	100.0	24	15	US-10-144-929-142	Sequence 142, App
18	20	100.0	25	11	US-09-983-802-624	Sequence 624, App
19	20	100.0	30	9	US-09-739-907-153	Sequence 153, App
20	20	100.0	31	15	US-10-023-282-882	Sequence 882, App
21	20	100.0	35	11	US-09-809-391-619	Sequence 619, App
22	20	100.0	36	9	US-09-764-869-670	Sequence 670, App
23	20	100.0	36	15	US-10-091-504-670	Sequence 670, App
24	20	100.0	39	15	US-10-106-698-7459	Sequence 7459, App
25	20	100.0	41	15	US-10-059-585-25	Sequence 25, App
26	20	100.0	44	9	US-09-764-869-894	Sequence 894, App
27	20	100.0	44	15	US-10-091-504-894	Sequence 894, App
28	20	100.0	45	9	US-09-864-761-43127	Sequence 43127, A
29	20	100.0	46	12	US-10-097-111-450	Sequence 450, App
30	20	100.0	48	10	US-09-952-4328-10	Sequence 10, App
31	20	100.0	48	15	US-10-012-542-143	Sequence 143, App
32	20	100.0	49	10	US-09-764-847-850	Sequence 850, App
33	20	100.0	49	15	US-10-092-154-850	Sequence 850, App
34	20	100.0	50	9	US-09-864-761-34130	Sequence 34120, A
35	20	100.0	50	15	US-10-023-282-431	Sequence 431, App
36	20	100.0	51	9	US-09-864-761-37683	Sequence 37683, A
37	20	100.0	51	9	US-09-864-761-42675	Sequence 42675, A
38	20	100.0	51	9	US-09-867-550-874	Sequence 874, App
39	20	100.0	52	9	US-09-864-761-43128	Sequence 43128, A
40	20	100.0	52	10	US-09-730-763-25	Sequence 25, App
41	20	100.0	53	15	US-10-106-698-4492	Sequence 4492, App
42	20	100.0	54	14	US-10-001-870-151	Sequence 151, App
43	20	100.0	55	9	US-09-764-860-306	Sequence 306, App
44	20	100.0	55	11	US-09-874-472-14	Sequence 14, App
45	20	100.0	55	11	US-09-874-472-16	Sequence 16, App

ALIGNMENTS

```
RESULT 1
US-09-300-425B-30
: Sequence 30, Application US/09300425B
: Publication No. US20030045681A1
: GENERAL INFORMATION:
: APPLICANT: NERI, Dario
: APPLICANT: TARLI, Lorenzo
: APPLICANT: VITI, Francesca
: APPLICANT: BIRCHER, Manfred
: TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
: TITLE OF INVENTION: CLAIMING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
: TITLE OF INVENTION: ANGIOGENESIS
: FILE REFERENCE: SCH-1733P1
: CURRENT APPLICATION NUMBER: US/09/300,425B
: CURRENT FILING DATE: 1999-04-28
: PRIOR APPLICATION NUMBER: 09/075,338
: PRIOR FILING DATE: 1998-05-11
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 30
: LENGTH: 4
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
: OTHER INFORMATION: antibody clone
US-09-300-425B-30
Query Match 100.0%; Score 20; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SFSF 4
1111
Db 1 SFSF 4

RESULT 2

US-09-947-387-131
; Sequence 131, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Molecule
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 122
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-947-387-131

Query Match 100.0%; Score 20; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 1 SFSF 4

RESULT 3
US-09-947-387-122
; Sequence 122, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Molecule
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 122
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-947-387-122

OTHER INFORMATION: Peptide
US-09-947-387-122

Query Match 100.0%; Score 20; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 1 SFSF 4

RESULT 4
US-09-947-387-33
; Sequence 33, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Molecule
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 33
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-947-387-33

Query Match 100.0%; Score 20; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 1 SFSF 4

RESULT 5
US-09-205-658-293
; Sequence 293, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15

NUMBER OF SEQ ID NOS: 328
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 293
LENGTH: 11
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-205-658-293

Query Match 100.0%; Score 20; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSE 4
1111
DB 8 SFSE 11

RESULT 6
US-09-934-251A-1
Sequence 1, Application US/09934251A
Patent No. US20020132017A1
GENERAL INFORMATION:
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
FILE REFERENCE: 108236.136
CURRENT APPLICATION NUMBER: US/09/934,251A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/368,607
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: US 08/762,537
PRIOR FILING DATE: 1996-12-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide chain of pylartin protein
US-09-934-251A-1

Query Match 100.0%; Score 20; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSE 4
1111
DB 5 SFSE 8

RESULT 7
US-09-879-957-187
Sequence 187, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
HOFFMAN, NO. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 187:
US-09-879-957-187

Query Match 100.0%; Score 20; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSE 4
1111
DB 1 SFSE 4

RESULT 8
US-09-879-957-202
Sequence 202, Application US/09879957
Patent No. US20020034755A1
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
HOFFMAN, NO. US20020034755A1h
KAY, Brian K.
FOWLES, Dana M.
MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 202:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Other
OTHER INFORMATION: May or may not have carboxy-terminal
amide and/or biotinylated N-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 202:
US-09-879-957-202

Query Match 100.0%; Score 20; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
||||
Db 1 SFSF 4

RESULT 9
US-09-934-251A-2
Sequence 2, Application US/09934251A
Patent No. US20020132017A1
GENERAL INFORMATION:
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
FILE REFERENCE: 108236.136
CURRENT APPLICATION NUMBER: US/09/934,251A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/368,607
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: US 08/762,537
PRIOR FILING DATE: 1996-12-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

Query Match 100.0%; Score 20; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
||||
Db 5 SFSF 8

RESULT 10
US-09-775-052-46
Sequence 46, Application US/09775052
Publication No. US20030054000A1
GENERAL INFORMATION:
APPLICANT: Dowdy, Steven F.
TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
FILE REFERENCE: 48861/1742
CURRENT APPLICATION NUMBER: US/09/775,052
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 09/208,966
PRIOR FILING DATE: 1998-12-10

PRIOR APPLICATION NUMBER: 60/082,402
PRIOR FILING DATE: 1998-04-20
PRIOR APPLICATION NUMBER: 60/069,012
PRIOR FILING DATE: 1997-12-10
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 46
LENGTH: 14
TYPE: PRT
ORGANISM: human
US-09-775-052-46

Query Match 100.0%; Score 20; DB 11; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
||||
Db 4 SFSF 7

RESULT 11
US-09-892-877-413
Sequence 413, Application US/09892877
Publication No. US2003007809A1
GENERAL INFORMATION:
APPLICANT: Ruben et. al.
TITLE OF INVENTION: 97 Human secreted proteins
FILE REFERENCE: P2028P1
CURRENT APPLICATION NUMBER: US/09/892,877
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 461
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 413
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-09-892-877-413

Query Match 100.0%; Score 20; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
||||
Db 11 SFSF 14

RESULT 12
US-09-948-783-413
Sequence 413, Application US/09948783
Publication No. US20030100051A1
GENERAL INFORMATION:
APPLICANT: Ruben et. al.
TITLE OF INVENTION: 97 Human secreted proteins
FILE REFERENCE: P2028P2
CURRENT APPLICATION NUMBER: US/09/948,783
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/231,846
PRIOR FILING DATE: 2000-09-11
PRIOR APPLICATION NUMBER: 09/892,877
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 09/437,658
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: PCT/US99/09847
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 60/085,093
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,094
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,105

PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,180
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,927
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,906
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,924
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,928
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,920
PRIOR FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 465
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 413
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-09-948-783-413

Query Match 100.0%; Score 20; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 11 SFSF 14

RESULT 13
US-10-260-212-3
Sequence 3, Application US/10260212
Publication No. US2003013180A1
GENERAL INFORMATION:
APPLICANT: SOCIETE DES PRODUITS NESTLE S.A.
TITLE OF INVENTION: Coffee Mannanase
FILE REFERENCE: 88265-6783
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: PCT/EP01/01549
PRIOR FILING DATE: 2001-02-13
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Coffea arabica
US-10-260-212-3

Query Match 100.0%; Score 20; DB 16; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 1 SFSF 4

RESULT 14
US-09-983-802-158
Sequence 158, Application US/09983802
Publication No. US20030022185A1
GENERAL INFORMATION:
APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/983,802
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357
PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,803
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,785
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,664
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672

Search completed: August 20, 2003, 13:16:46
Job time : 5.33735 secs

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 158
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (23)
OTHER INFORMATION: Xaa equals stop translation
US-09-983-802-158

Query Match 100.0%; Score 20; DB 11; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
1111
DB 17 SFSF 20

RESULT 15
US-10-097-065-345
Sequence 345, Application US/10097065
Publication No. US20030055236A1
GENERAL INFORMATION:
APPLICANT: Moore, Paul A. et al.
TITLE OF INVENTION: 110 Human Secreted Proteins
FILE REFERENCE: P2021P1
CURRENT APPLICATION NUMBER: US/10/097,065
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: PCT/US98/27059
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,007
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,057
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,006
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,369
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,367
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,368
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,169
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,008
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,365
PRIOR FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 345
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-065-345

Query Match 100.0%; Score 20; DB 15; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
1111
DB 17 SFSF 14

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:35:06 : Search time 40.9398 Seconds
(without alignments)
85.038 Million cell updates/sec

Title: US-09-512-082-30
Perfect score: 20
Sequence: 1 SFSF 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents, AA Main:*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	US-09-075-338C-30	Sequence 30, Appl
2	20	100.0	4	US-09-300-425B-30	Sequence 30, Appl

3	20	100.0	4	19	US-09-512-082-30	Sequence 30, Appl
4	20	100.0	5	1	PCT-US98-21231-131	Sequence 131, App
5	20	100.0	5	1	PCT-US99-16423-131	Sequence 131, App
6	20	100.0	5	19	US-09-583-225-131	Sequence 131, App
7	20	100.0	5	24	US-09-947-387-131	Sequence 131, App
8	20	100.0	5	27	US-10-138-375-131	Sequence 122, App
9	20	100.0	6	1	PCT-US98-21231-122	Sequence 122, App
10	20	100.0	6	1	PCT-US99-16423-122	Sequence 122, App
11	20	100.0	6	19	US-09-583-225-122	Sequence 122, App
12	20	100.0	6	24	US-09-947-387-122	Sequence 122, App
13	20	100.0	6	27	US-10-138-375-122	Sequence 122, App
14	20	100.0	7	1	PCT-US98-21231-133	Sequence 33, Appl
15	20	100.0	7	1	PCT-US99-16423-133	Sequence 33, Appl
16	20	100.0	7	19	US-09-583-225-133	Sequence 33, Appl
17	20	100.0	7	24	US-09-947-387-133	Sequence 33, Appl
18	20	100.0	7	27	US-10-138-375-133	Sequence 33, Appl
19	20	100.0	7	30	US-10-405-027-851	Sequence 5551, App
20	20	100.0	8	18	US-09-412-863-5521	Sequence 6521, App
21	20	100.0	8	18	US-09-412-863-5525	Sequence 6525, App
22	20	100.0	8	18	US-09-412-863A-6521	Sequence 6521, App
23	20	100.0	8	18	US-09-412-863A-6525	Sequence 6525, App
24	20	100.0	9	18	US-09-412-863A-4572	Sequence 4572, App
25	20	100.0	9	18	US-09-412-863-12694	Sequence 12694, A
26	20	100.0	9	18	US-09-412-863-13945	Sequence 13945, A
27	20	100.0	9	18	US-09-412-863A-4422	Sequence 4422, App
28	20	100.0	9	18	US-09-412-863A-12694	Sequence 12694, A
29	20	100.0	9	18	US-09-412-863A-13945	Sequence 13945, A
30	20	100.0	9	20	US-09-647-372-46	Sequence 46, Appl
31	20	100.0	9	20	US-09-647-372B-46	Sequence 46, Appl
32	20	100.0	10	18	US-09-412-863-4572	Sequence 4572, App
33	20	100.0	10	18	US-09-412-863-6522	Sequence 6522, App
34	20	100.0	10	18	US-09-412-863-6526	Sequence 6526, App
35	20	100.0	10	18	US-09-412-863-12693	Sequence 12693, A
36	20	100.0	10	18	US-09-412-863A-4572	Sequence 4572, App
37	20	100.0	10	18	US-09-412-863A-6522	Sequence 6522, App
38	20	100.0	10	18	US-09-412-863A-6526	Sequence 6526, App
39	20	100.0	10	18	US-09-412-863A-12693	Sequence 12693, A
40	20	100.0	11	1	PCT-US99-28529-293	Sequence 293, App
41	20	100.0	11	16	US-09-205-658-293	Sequence 293, App
42	20	100.0	11	16	US-09-412-863-318	Sequence 318, App
43	20	100.0	11	18	US-09-412-863-323	Sequence 323, App
44	20	100.0	11	18	US-09-412-863-6523	Sequence 6523, App
45	20	100.0	11	18	US-09-412-863-6524	Sequence 6524, App

ALIGNMENTS

RESULT 1
US-09-075-338C-30
Sequence 30, Application US/09075338C
GENERAL INFORMATION:
APPLICANT: NERI, Danilo
APPLICANT: TARLI, Lorenzo
APPLICANT: VITI, Francesca
APPLICANT: BIRCHIERI, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
FILE REFERENCE: SCH-1733
CURRENT APPLICATION NUMBER: US/09/075,338C
CURRENT FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
OTHER INFORMATION: antibody clone
US-09-075-338C-30

Query Match 100.0%; Score 20; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SFSF 4
|||||
Db 1 SFSF 4

RESULT 2

US-09-300-425B-30
; Sequence 30, Application US/09300425B
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-300-425B-30

Query Match Best Local Similarity 100.0%; Score 20; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SFSF 4
|||||
Db 1 SFSF 4

RESULT 3

US-09-512-082-30
; Sequence 30, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-512-082-30

Query Match 100.0%; Score 20; DB 19; Length 4;

Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SFSF 4
|||||
Db 1 SFSF 4

RESULT 4

PCT-US98-21231-131
; Sequence 131, Application PC/TUS9821231
; GENERAL INFORMATION:
; APPLICANT: Cytovia, Inc.
; TITLE OF INVENTION: Novel Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.029PC02
; CURRENT APPLICATION NUMBER: PCT/US98/21231
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 131
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
PCT-US98-21231-131

Query Match Best Local Similarity 100.0%; Score 20; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SFSF 4
|||||
Db 1 SFSF 4

RESULT 5

PCT-US99-16423-131
; Sequence 131, Application PC/TUS9916423
; GENERAL INFORMATION:
; APPLICANT: Cytovia, Inc.
; APPLICANT: Zhang, Han-zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Protea
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.003PC01
; CURRENT APPLICATION NUMBER: PCT/US99/16423
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US 60/093,642
; EARLIER FILING DATE: 21-JUL-1998
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 131
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
PCT-US99-16423-131

Query Match 100.0%; Score 20; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 1 SFSF 4

RESULT 6

US-09-583-225-131
; Sequence 131, Application US/09583225
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Protease
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/09/583,225
; PRIOR FILING DATE: 2000-05-30
; PRIOR FILING DATE: 09/357,952
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-583-225-131

Query Match 100.0%; Score 20; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 1 SFSF 4

RESULT 7

US-09-947-387-131
; Sequence 131, Application US/09947387
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: Novel Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-947-387-131

Query Match 100.0%; Score 20; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 1 SFSF 4

RESULT 8

US-10-138-375-131
; Sequence 131, Application US/10138375
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Prote
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/10/138,375
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/357,952
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: US 60/093,642
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-138-375-131

Query Match 100.0%; Score 20; DB 27; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 1 SFSF 4

RESULT 9

PCT-US98-21231-122
; Sequence 122, Application PC/YUS9821231
; GENERAL INFORMATION:
; APPLICANT: Cytovia, Inc.
; TITLE OF INVENTION: Novel Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.029PC02
; CURRENT APPLICATION NUMBER: PCT/US98/21231
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
PCT-US98-21231-122

Query Match 100.0%; Score 20; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 10
PCT-US99-16423-122
Sequence 122, Application PC/TUS9916423
GENERAL INFORMATION:
APPLICANT: Cytovia, Inc.
APPLICANT: Zhang, Han-Zhong
APPLICANT: Cai, Sui Xiong
APPLICANT: Drewe, John A.
APPLICANT: Yang, Wu
TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Proteases
FILE REFERENCE: 1735.003PC01
CURRENT APPLICATION NUMBER: PCT/US99/16423
PRIOR FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US 60/093,642
NUMBER OF SEQ ID NOS: 139
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 122
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
PCT-US99-16423-122

Query Match 100.0%; Score 20; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 11
US-09-583-225-122
Sequence 122, Application US/09583225
GENERAL INFORMATION:
APPLICANT: Zhang, Han-Zhong
APPLICANT: Cai, Sui Xiong
APPLICANT: Drewe, John A.
APPLICANT: Yang, Wu
TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Protease
FILE REFERENCE: 1735.0030001
CURRENT APPLICATION NUMBER: US/09/583,225
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 09/357,952
PRIOR FILING DATE: 21-JUL-1999
NUMBER OF SEQ ID NOS: 139
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 122
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-583-225-122

Query Match 100.0%; Score 20; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 12
US-09-947-387-122
Sequence 122, Application US/09947387
GENERAL INFORMATION:
APPLICANT: Weber, Eckard
APPLICANT: Cai, Sui Xiong
APPLICANT: Keana, John F.W.
APPLICANT: Drewe, John A.
APPLICANT: Zhang, Han-Zhong
TITLE OF INVENTION: Novel Fluorescent or Fluorescent Reporter Molecules and
TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
FILE REFERENCE: 1735.0250003
CURRENT APPLICATION NUMBER: US/09/947,387
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/061,582
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: US 60/145,746
PRIOR FILING DATE: 1998-03-03
PRIOR APPLICATION NUMBER: US 09/168,888
PRIOR FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 142
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 122
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-947-387-122

Query Match 100.0%; Score 20; DB 24; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
DB 1 SF5F 4

RESULT 13
US-10-138-375-122
Sequence 122, Application US/10138375
GENERAL INFORMATION:
APPLICANT: Zhang, Han-Zhong
APPLICANT: Cai, Sui Xiong
APPLICANT: Drewe, John A.
APPLICANT: Yang, Wu
TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Prote
FILE REFERENCE: 1735.0030001
CURRENT APPLICATION NUMBER: US/10/138,375
PRIOR FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/357,952
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 1999-07-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/093,642

PRIOR FILING DATE: EARLIER FILING DATE: 21-JUL-1998
 ; NUMBER OF SEQ ID NOS: 139
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 122
 ; LENGTH: 6
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; US-10-138-375-122

Query Match 100.0%; Score 20; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.2e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
 ||||
 Db 1 SFSF 4

RESULT 14
 PCT-US98-21231-33
 ; Sequence 33, Application PC/TUS9821231
 ; GENERAL INFORMATION:
 ; APPLICANT: Cytovia, Inc.
 ; TITLE OF INVENTION: Novel Fluorogenic or Fluorescent Reporter Molecules and
 ; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
 ; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
 ; TITLE OF INVENTION: Use Thereof
 ; FILE REFERENCE: 1735.029PC02
 ; CURRENT APPLICATION NUMBER: PCT/US98/21231
 ; CURRENT FILING DATE: 1998-10-09
 ; EARLIER APPLICATION NUMBER: US 60/061,582
 ; EARLIER FILING DATE: 1997-10-10
 ; EARLIER APPLICATION NUMBER: US 09/033,661
 ; EARLIER FILING DATE: 1998-03-03
 ; NUMBER OF SEQ ID NOS: 142
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 33
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Peptide
 PCT-US98-21231-33

Query Match 100.0%; Score 20; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.2e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
 ||||
 Db 1 SFSF 4

RESULT 15
 PCT-US99-16423-33
 ; Sequence 33, Application PC/TUS9916423
 ; GENERAL INFORMATION:
 ; APPLICANT: Cytovia, Inc.
 ; APPLICANT: Zhang, Han-Zhong
 ; APPLICANT: Cai, Sui Xiong
 ; APPLICANT: Yang, Wu
 ; APPLICANT: Drewe, John A.
 ; TITLE OF INVENTION: Novel Fluorescence Dyes and Their Applications for Whole-Cell
 ; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Proteases
 ; TITLE OF INVENTION: Other Enzymes and the Use Thereof
 ; FILE REFERENCE: 1735.003PC01
 ; CURRENT APPLICATION NUMBER: PCT/US99/16423
 ; CURRENT FILING DATE: 1999-07-21
 ; EARLIER APPLICATION NUMBER: US 60/093,642

EARLIER FILING DATE: 21-JUL-1998
 ; NUMBER OF SEQ ID NOS: 139
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 33
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Peptide
 PCT-US99-16423-33

Query Match 100.0%; Score 20; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.2e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
 ||||
 Db 1 SFSF 4

Search completed: August 20, 2003, 13:13:40
 Job time : 41.9398 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 1.15663 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-30
Perfect score: 20
Sequence: 1 SFSF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues
Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCR_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	11	5	US-09-963-693A-293
2	20	100.0	26	6	US-10-389-674-88
3	20	100.0	52	6	US-10-368-433A-23
4	20	100.0	52	6	US-10-631-402-3325
5	20	100.0	52	6	US-10-631-441-3325
6	20	100.0	58	6	US-10-631-402-3330
7	20	100.0	58	6	US-10-631-441-3330
8	20	100.0	61	6	US-10-603-113-24077
9	20	100.0	61	6	US-10-603-113-23238
10	20	100.0	61	6	US-10-603-113-24772
11	20	100.0	61	6	US-10-603-113-26821
12	20	100.0	62	6	US-10-291-128-150
13	20	100.0	62	6	US-10-291-128-151
14	20	100.0	62	6	US-10-603-113-21539
15	20	100.0	62	6	US-10-603-113-23627
16	20	100.0	64	6	US-10-603-113-24631
17	20	100.0	64	6	US-10-603-113-26647
18	20	100.0	65	6	US-10-603-113-25269
19	20	100.0	65	6	US-10-612-783-6240
20	20	100.0	66	6	US-10-617-320-2625
21	20	100.0	66	6	US-10-603-113-22497
22	20	100.0	66	6	US-10-603-113-26882
23	20	100.0	67	6	US-10-603-113-24997
24	20	100.0	67	6	US-10-603-113-26780
25	20	100.0	68	1	PCR-US02-41612A-489
26	20	100.0	69	6	US-10-291-265-430

27	20	100.0	69	6	US-10-291-265-902	Sequence 902, App
28	20	100.0	70	6	US-10-603-113-22184	Sequence 22184, A
29	20	100.0	70	6	US-10-603-113-25419	Sequence 25419, A
30	20	100.0	71	6	US-10-273-573-8204	Sequence 8204, Ap
31	20	100.0	71	6	US-10-617-320-4664	Sequence 4664, Ap
32	20	100.0	71	6	US-10-603-113-23802	Sequence 23802, A
33	20	100.0	72	6	US-10-603-113-25759	Sequence 25759, A
34	20	100.0	72	6	US-10-612-783-5219	Sequence 5219, Ap
35	20	100.0	73	6	US-10-603-113-21888	Sequence 21888, A
36	20	100.0	73	6	US-10-603-113-27444	Sequence 27444, A
37	20	100.0	74	6	US-10-617-320-4732	Sequence 4732, Ap
38	20	100.0	74	6	US-10-603-113-24812	Sequence 24812, A
39	20	100.0	74	6	US-10-603-113-26215	Sequence 26215, A
40	20	100.0	74	6	US-10-631-402-1780	Sequence 1780, Ap
41	20	100.0	74	6	US-10-631-441-1780	Sequence 1780, Ap
42	20	100.0	75	6	US-10-603-113-21184	Sequence 21184, A
43	20	100.0	77	6	US-10-603-113-24293	Sequence 24293, A
44	20	100.0	77	6	US-10-603-113-25632	Sequence 25632, A
45	20	100.0	79	6	US-10-603-113-24061	Sequence 24061, A

ALIGNMENTS

```

RESULT 1
US-09-963-693A-293
; Sequence 293, Application US/09963693A
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Osg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/963, 693A
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 09/205, 658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857, 076
; PRIOR FILING DATE: 1997-05-15
; PRIOR APPLICATION NUMBER: 08/888, 534
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: US98/10080
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 293
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-963-693A-293
Query Match          100.0%; Score 20; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 SFSF 4
Db      8 SFSF 11
RESULT 2
US-10-389-674-88
; Sequence 88, Application US/10389674
; GENERAL INFORMATION:
; APPLICANT: HEINRICH, VOLKER
; APPLICANT: CHEN, TEDDY
; APPLICANT: PATEN, PHILLIP A.
; TITLE OF INVENTION: IPR-ALPHA HOMOLOGUES
; FILE REFERENCE: 02-101510/0140.002
; CURRENT APPLICATION NUMBER: US/10/389, 674
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US/09/685, 189
; PRIOR FILING DATE: 2000-10-06

```

```
;; PRIOR APPLICATION NUMBER: 09/415,183
;; PRIOR FILING DATE: 1999-10-07
;; NUMBER OF SEQ ID NOS: 88
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 88
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic amino acid
US-10-389-674-88
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SFSF 4
    ||||
Db 11 SFSF 14
```

```
RESULT 3
US-10-368-433A-23
; Sequence 23, Application US/10368433A
; GENERAL INFORMATION:
; APPLICANT: JACKSON, MARY
; APPLICANT: GIGOUDEL, BRIGITTE
; TITLE OF INVENTION: DESATURASE ANTIGEN OF MYCOBACTERIUM TUBERCULOSIS
; FILE REFERENCE: 05394.0009-01
; CURRENT APPLICATION NUMBER: US/10/368,433A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: PCT/IB97/009923
; PRIOR FILING DATE: 1997-07-25
; PRIOR APPLICATION NUMBER: 60/022,713
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Pseudomonas sp.
US-10-368-433A-23
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SFSF 4
    ||||
Db 6 SFSF 9
```

```
RESULT 4
US-10-631-402-3325
; Sequence 3325, Application US/10631402
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,402
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
```

```
;; PRIOR APPLICATION NUMBER: US 08/905,279
;; PRIOR FILING DATE: 1997-08-01
;; PRIOR APPLICATION NUMBER: US 08/904,468
;; PRIOR FILING DATE: 1997-08-01
;; PRIOR APPLICATION NUMBER: US 08/905,134
;; PRIOR FILING DATE: 1997-08-01
;; PRIOR APPLICATION NUMBER: US 08/905,133
;; PRIOR FILING DATE: 1997-08-01
;; NUMBER OF SEQ ID NOS: 3475
;; SOFTWARE: Patent.pm
;; SEQ ID NO 3325
;; LENGTH: 52
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; OTHER INFORMATION: Substantia nigra
;; NAME/KEY: SIGNAL
;; LOCATION: -38..-1
;; OTHER INFORMATION: Von Heijne matrix
;; FEATURE:
;; OTHER INFORMATION: score 8.20
;; FEATURE:
;; OTHER INFORMATION: seq SFXFLALCASFS/FF
;; NAME/KEY: UNSURE
;; LOCATION: 4
;; OTHER INFORMATION: Xaa = Cys,Phe
;; NAME/KEY: UNSURE
;; LOCATION: -11
;; OTHER INFORMATION: Xaa = Leu,Tyr
US-10-631-402-3325
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SFSF 4
    ||||
Db 36 SFSF 39
```

```
RESULT 5
US-10-631-441-3325
; Sequence 3325, Application US/10631441
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,441
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
```

```

; SOFTWARE: Patent.pm
; SEQ ID NO 3325
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
;   OTHER INFORMATION: Substantia nigra
; NAME/KEY: SIGNAL
; LOCATION: -38...-1
; OTHER INFORMATION: Von Heijne matrix
; FEATURE:
;   OTHER INFORMATION: score 8.20
; FEATURE:
;   OTHER INFORMATION: seq SFXFLALCASRS/FF
; NAME/KEY: UNSURE
; LOCATION: 4
; OTHER INFORMATION: Xaa - Cys,Phe
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: -11
; OTHER INFORMATION: Xaa - Leu,Tyr
US-10-631-441-3325

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SFSF 4
        ||||
Db       36 SFSF 39

RESULT 6
US-10-631-402-3330
; Sequence 3330, Application US/10631402
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,402
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 3330
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
;   OTHER INFORMATION: Cerebellum
; NAME/KEY: SIGNAL
```

```

; LOCATION: -47...-1
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 7.60
; OTHER INFORMATION: seq CLATLTLFHTSFS/FQ
US-10-631-402-3330

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SFSF 4
        ||||
Db       45 SFSF 48

RESULT 7
US-10-631-441-3330
; Sequence 3330, Application US/10631441
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,441
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 3330
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
;   OTHER INFORMATION: Cerebellum
; NAME/KEY: SIGNAL
; LOCATION: -47...-1
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 7.60
; OTHER INFORMATION: seq CLATLTLFHTSFS/FQ
US-10-631-441-3330

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SFSF 4
        ||||
Db       45 SFSF 48

RESULT 8
US-10-603-113-24077
; Sequence 24077, Application US/10603113
; GENERAL INFORMATION:
```

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603,113
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/248,796
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO 24077
LENGTH: 60
TYPE: PRT
ORGANISM: Candida albicans
US-10-603-113-24077

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 60;
Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
||||
DB 41 SF5F 44

RESULT 9
US-10-603-113-23238
Sequence 23238, Application US/10603113
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603,113
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/248,796
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO 23238
LENGTH: 61
TYPE: PRT
ORGANISM: Candida albicans
US-10-603-113-23238

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 61;
Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
||||
DB 31 SF5F 34

RESULT 10
US-10-603-113-24772
Sequence 24772, Application US/10603113
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603,113
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/248,796
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO 24772
LENGTH: 61
TYPE: PRT
ORGANISM: Candida albicans
NAME/KEY: UNSURE
LOCATION: (?)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-10-603-113-24772

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 61;
Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
||||
DB 57 SF5F 60

RESULT 11
US-10-603-113-26821
Sequence 26821, Application US/10603113
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/10/603,113
CURRENT FILING DATE: 2003-06-24
PRIOR APPLICATION NUMBER: US/09/248,796
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 28206
SEQ ID NO 26821
LENGTH: 61
TYPE: PRT
ORGANISM: Candida albicans
US-10-603-113-26821

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 61;
Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SF5F 4
||||
DB 30 SF5F 33

RESULT 12
US-10-291-128-150
Sequence 150, Application US/10291128
GENERAL INFORMATION:
APPLICANT: Nuveto, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP4
CURRENT APPLICATION NUMBER: US/10/291,128
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/10484
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/668,680
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/695,618
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 09/728,711
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/808,701
PRIOR FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 172
SOFTWARE: Custom
SEQ ID NO 150
LENGTH: 62
TYPE: PRT
ORGANISM: Homo sapiens
US-10-291-128-150

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 62;
Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 1 SFSF 4

RESULT 13
US-10-291-128-151
; Sequence 151, Application US/10291128
; GENERAL INFORMATION:
; APPLICANT: Nuvelo, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP4
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/10484
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/666,660
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/808,701
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO: 151
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-128-151

Query Match 100.0%; Score 20; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 1 SFSF 4

RESULT 14
US-10-603-113-21539
; Sequence 21539, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Welnstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO: 21539
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-21539

Query Match 100.0%; Score 20; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 58 SFSF 61

RESULT 15
US-10-603-113-23627
; Sequence 23627, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Welnstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO: 23627
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-23627

Query Match 100.0%; Score 20; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
||||
DB 13 SFSF 16

Search completed: August 20, 2003, 12:45:13
Job time: 1.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 2.79518 Seconds
(without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-30

Perfect score: 20

Sequence: 1 SFSF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	10	2 S38305	lectin GNL2 alpha
2	20	100.0	24	2 S08292	lectin - smooth ra
3	20	100.0	32	2 A03367	lectin - Macrotylo
4	20	100.0	40	2 A48431	agglutinin, EVA (N
5	20	100.0	42	2 E84033	hypothetical prote
6	20	100.0	50	2 H82423	hypothetical prote
7	20	100.0	50	2 T07306	hypothetical prote
8	20	100.0	50	2 G87539	hypothetical prote
9	20	100.0	51	2 G81188	hypothetical prote
10	20	100.0	54	1 WRBP66	early protein gp6
11	20	100.0	54	1 WRBP66	early protein gp6
12	20	100.0	54	1 JN0034	early protein gp6
13	20	100.0	57	4 TMBP11	hypothetical immun
14	20	100.0	60	2 T07328	hypothetical prote
15	20	100.0	61	2 D84839	hypothetical prote
16	20	100.0	63	2 S12241	hypothetical prote
17	20	100.0	63	2 E82584	hypothetical prote
18	20	100.0	64	2 D81095	hypothetical prote
19	20	100.0	66	2 S45377	hypothetical prote
20	20	100.0	66	2 AF3083	hypothetical prote
21	20	100.0	67	2 G72372	hypothetical prote
22	20	100.0	70	2 T12945	hypothetical prote
23	20	100.0	71	2 E95109	conserved hypotnet
24	20	100.0	71	2 G83792	hypothetical prote
25	20	100.0	71	2 H97977	conserved hypotnet
26	20	100.0	72	2 I40354	hypothetical prote
27	20	100.0	73	2 T07335	hypothetical prote
28	20	100.0	75	2 E81252	hypothetical prote
29	20	100.0	76	2 B97816	hypothetical prote

30	20	100.0	78	2 D82182	hypothetical prote
31	20	100.0	79	2 T42000	hypothetical prote
32	20	100.0	80	2 S26991	hypothetical prote
33	20	100.0	83	2 C47188	MHC class II histo
34	20	100.0	83	2 T17809	hypothetical prote
35	20	100.0	87	2 G90954	hypothetical prote
36	20	100.0	87	2 D85803	hypothetical prote
37	20	100.0	89	2 C42525	A-ORF-P protein -
38	20	100.0	89	2 J01822	SalPa protein - va
39	20	100.0	91	2 B96574	hypothetical prote
40	20	100.0	91	2 T06963	hypothetical prote
41	20	100.0	93	2 AD1247	hypothetical prote
42	20	100.0	93	2 AH1609	hypothetical prote
43	20	100.0	94	2 AE3247	hypothetical prote
44	20	100.0	96	2 T29674	hypothetical prote
45	20	100.0	97	2 C75415	hypothetical prote

ALIGNMENTS

RESULT 1

S38305
lectin GNL2 alpha chain - kidney bean (fragment)
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
C:Accession: S38305
R:Kamemura, K.; Furutachi, Y.; Umekawa, H.; Takahashi, T.
Biochim. Biophys. Acta 1158, 181-188, 1993
A:Title: Purification and characterization of novel lectins from Great Northern bean,
A:Reference number: S38304; M01D:94002183; PMID:839319
A:Accession: S38305
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <RMM>

Query Match 100.0%; Score 20; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 5 SFSF 8

RESULT 2

S08292
lectin - smooth rattlebox (fragment)
C:Species: Crotalaria pallida (smooth rattlebox)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1993
C:Accession: S08292
R:Nguyen, O.K.; Guillaume, J.L.; Hoebeke, J.
Biochim. Biophys. Acta 1033, 210-213, 1990
A:Title: A blood group A specific lectin from the seeds of Crotalaria striata.
A:Reference number: S08292; M01D:90167102; PMID:2306467
A:Accession: S08292
A:Molecule type: protein
A:Residues: 1-24 <NGU>
A:Keywords: the source is designated as Crotalaria striata

Query Match 100.0%; Score 20; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 5 SFSF 8

RESULT 3

A03367
lectin - Macrotyloma axillare (fragment)

C:Species: Macrotyloma axillare
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1993
 C:Accession: A03367
 R:Haylett, T.; Swart, L.S.
 S. Afr. J. Chem. 35, 33-36, 1982
 A:Reference number: A03367
 A:Accession: A03367
 A:Molecule type: protein
 A:Residues: 1-32 <HAY>
 C:Comment: This lectin from legume seeds is a tetramer of noncovalently bound identical
 e carbohydrates galactose, glucosamine, mannose, and fucose. It agglutinates erythrocyte
 C:Comment: This fragment is related to the amino end of the lentil lectin beta chain.
 C:Comment: Dolichos axillaris (Macrotyloma axillare) is an African legume, also now grown
 C:Superfamily: plant lectin

Query Match 100.0%; Score 20; DB 2; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 |||||
 DB 5 SFSF 8

RESULT 4
 A48431
 agglutinin, EVA (N-terminal) - Erythrina variegata (fragment)
 C:Species: Erythrina variegata
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 18-Sep-1998
 C:Accession: A48431
 R:Li, H.; Yamamoto, K.; Kawashima, H.; Osawa, T.
 Glycoconj. J. 7, 311-322, 1990.
 A:Title: Structural requirements for the binding of oligosaccharides to immobilized lect
 A:Reference number: A48431; MUID:92361170; PMID:2136348
 A:Contents: var. orientalis, seeds
 A:Accession: A48431
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-40 <LI1>
 A:Note: sequence extracted from NCBI backbone (NCBIP:110983)
 C:Superfamily: plant lectin

Query Match 100.0%; Score 20; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 |||||
 DB 5 SFSF 8

RESULT 5
 E84033
 hypothetical protein BH3069 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: E84033
 R:Takami, H.; Nakasone, K.; Takai, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: E84033
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-42 <STO>
 A:Cross-references: GB:AP001517; GB:BA000004; NID:g10175500; PIDN:BA06788.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3069

Query Match 100.0%; Score 20; DB 2; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
 |||||
 DB 35 SFSF 38

RESULT 6
 H82423
 hypothetical protein VCA0733 [imported] - Vibrio cholerae (strain N16961 serogroup O1
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: H82423
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
 Chardson, D.; Ermiolen, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: H82423
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-50 <HE1>
 A:Cross-references: GB:AE004402; GB:AE003853; NID:g9658148; PIDN:AAF96632.1; GSPDB:GN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0733
 A:Map position: 2

Query Match 100.0%; Score 20; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 |||||
 DB 31 SFSF 34

RESULT 7
 T07306
 hypothetical protein 50 - Chlorella vulgaris chloroplast
 C:Species: chloroplast Chlorella vulgaris
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
 C:Accession: T07306
 R:Wakusugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na
 Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
 A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C
 A:Reference number: 215985; MUID:97303241; PMID:9155184
 A:Accession: T07306
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-50 <WAK>
 A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BA057954.1; PID:g2224470
 C:Genetics:
 A:Genome: chloroplast
 C:Keywords: chloroplast

Query Match 100.0%; Score 20; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFSF 4
 |||||
 DB 45 SFSF 48

RESULT 8
 G97539
 hypothetical protein AGR_C_2731 [imported] - Agrobacterium tumefaciens (strain C58, C
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: G97539
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,

Science 294..2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97539
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-50 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87272.1; PID:915156562; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_2731
A:Map position: circular chromosome

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 6 SFSF 9

RESULT 9
G81188
hypothetical protein NMB0517 [imported] - *Neisseria meningitidis* (strain MC58 serogroup
C:Species: *Neisseria meningitidis*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81188
R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grand, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A>Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: G81188
A:Molecule type: DNA
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <TET>
A:Cross-references: GB:AE002408; GB:AE002098; NID:g7225740; PIDN:AAF40949.1; PID:g722574
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0517

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 45 SFSF 48

RESULT 10
WBP66
early protein gp6 - phage PZA
C:Species: phage PZA
A>Note: host *Bacillus subtilis*
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 23-Jul-1999
C:Accession: B29004
R:Paces, V.; Vilek, C.; Urbanek, P.; Hostomsky, Z.
Gene 44, 115-120, 1986
A>Title: Nucleotide sequence of the right early region of *Bacillus subtilis* phage PZA-CC
A:Reference number: A91551; MUID:87031575; PMID:3095189
A:Accession: B29004
A:Molecule type: DNA
A:Residues: 1-54 <PAC>
A:Cross-references: GB:M1813; GB:M1904; GB:M13905; NID:g216046; PIDN:AAA88495.1; PID:g
C:Genetics:
A:Gene: 16.6
C:Superfamily: phage PZA early protein gp6
C:Keywords: early protein

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 54;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 12 SFSF 15

RESULT 11
WBP66
early protein gp6 - phage phi-15
C:Species: phage phi-15
A>Note: host *Bacillus subtilis*
C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 23-Jul-1999
C:Accession: J50196
R:Benes, V.; Arnold, L.; Smrt, J.; Paces, V.
Gene 75, 341-347, 1989
A>Title: Nucleotide sequence of the right early region of *Bacillus* phage phi-15 and c
A:Reference number: J50192; MUID:89232766; PMID:2497055
A:Accession: J50196
A:Molecule type: DNA
A:Residues: 1-54 <BEN>
A:Cross-references: GB:M28830; NID:g215438; PIDN:AAA32333.1; PID:g215443
C:Genetics:
A:Gene: 16.6
C:Superfamily: phage PZA early protein gp6
C:Keywords: early protein

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 12 SFSF 15

RESULT 12
JN0034
early protein gp6 - phage phi-29
N:Alternate names: ORF 16.6 protein
C:Species: phage phi-29
A>Note: host *Bacillus amyloliquefaciens*
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: JN0034
R:Garvey, K.J.; Yoshikawa, H.; Ito, J.
Gene 40, 301-309, 1985
A>Title: The complete sequence of the *Bacillus* phage phi-29 right early region.
A:Reference number: JN0030; MUID:86165872; PMID:3007295
A:Accession: JN0034
A:Molecule type: DNA
A:Residues: 1-54 <GAR>
A:Cross-references: GB:M14430; NID:g215321; PIDN:AAA88353.1; PID:g1196691
C:Superfamily: phage PZA early protein gp6

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
DB 12 SFSF 15

RESULT 13
JNBP11
hypothetical immunity region protein 11 - *Bacillus* phage phi-105
C:Species: *Bacillus* phage phi-105
A>Note: host *Bacillus subtilis*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: A27234
R:Cully, D.F.; Garro, A.J.
Gene 38, 153-164, 1985

A:Title: Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophage
A:Reference number: A91535; MUID:86056972; PMID:3934047
A:Accession: A27234
A:Molecule type: DNA
A:Residues: 1-57 <CUT>
A:Cross-references: GB:M1920; NID:g215477; PIDN:AAA8398.1; PID:g1196716
C:Comment: This is the hypothetical translation of a sequence that was not reported as a

Search completed: August 20, 2003, 12:42:20
Job time : 4.79518 secs

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
|||||
Db 6 SFSF 9

RESULT 14

T07328
hypothetical protein 60c - Chlorella vulgaris chloroplast
C:Species: chloroplast Chlorella vulgaris
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07328
R:Nakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chld
A:Reference number: Z15985; MUID:97303241; PMID:9159184
A:Accession: T07328
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-60 <MAK>
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57976.1; PID:g2224492
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 60;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
|||||
Db 39 SFSF 42

RESULT 15

D84839
hypothetical protein At2g41230 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84839
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Nuss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84839
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-61 <STO>
A:Cross-references: GB:AE002093; NID:g3894199; PIDN:AAC78548.1; GSPDB:GM00139
C:Genetics:
A:Gene: At2g41230
A:Map position: 2

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
|||||
Db 38 SFSF 41

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 1.51807 Seconds
(without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-30

Perfect score: 20
Sequence: 1 SF5F 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	24	1	LEC_CROST
2	20	100.0	32	1	LEC_DOLAX
3	20	100.0	42	1	PSAJ_NEPOL
4	20	100.0	54	1	VG66_BPPH2
5	20	100.0	54	1	VG66_BPPH5
6	20	100.0	54	1	VG66_BPPH2
7	20	100.0	57	1	YIMB_BPPH1
8	20	100.0	63	1	BP4A_BRANA
9	20	100.0	89	1	YVAP_VACCC
10	20	100.0	91	1	YCXF_CVABA
11	20	100.0	100	1	POL_HV113
12	20	100.0	103	1	Y067_METIA
13	20	100.0	110	1	Y123_METIA
14	20	100.0	111	1	YGC4_YEAST
15	20	100.0	113	1	Y465_AOUAE
16	20	100.0	114	1	CU05_HOMAM
17	20	100.0	116	1	Y457_YEAST
18	20	100.0	118	1	POL_HV1C4
19	20	100.0	120	1	GTRA_BPP22
20	20	100.0	120	1	GTRA_BPPS2
21	20	100.0	120	1	GTRA_BPPS5
22	20	100.0	120	1	GTRA_BPPS5
23	20	100.0	120	1	GTRA_BPPS5
24	20	100.0	121	1	GTRA_BPPS5
25	20	100.0	121	1	GTRA_BPPS5
26	20	100.0	135	1	Y067_METIA
27	20	100.0	136	1	Y067_METIA
28	20	100.0	141	1	Y067_METIA
29	20	100.0	143	1	Y067_METIA
30	20	100.0	143	1	Y067_METIA
31	20	100.0	145	1	Y067_METIA
32	20	100.0	146	1	Y067_METIA
33	20	100.0	147	1	Y067_METIA

34	20	100.0	147	1	Y067_METIA
35	20	100.0	150	1	Y067_METIA
36	20	100.0	154	1	Y067_METIA
37	20	100.0	160	1	Y067_METIA
38	20	100.0	161	1	Y067_METIA
39	20	100.0	163	1	Y067_METIA
40	20	100.0	164	1	Y067_METIA
41	20	100.0	164	1	Y067_METIA
42	20	100.0	170	1	Y067_METIA
43	20	100.0	185	1	Y067_METIA
44	20	100.0	185	1	Y067_METIA
45	20	100.0	186	1	Y067_METIA

ALIGNMENTS

RESULT 1	LEC_CROST	STANDARD;	PRT;	24 AA.
AC	P16351;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DE	01-MAY-1992 (Rel. 22, Last annotation update)			
DE	Lectin (Fragment).			
OS	Crotalaria striata (Smooth ratlebox) (Crotalaria pallida).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Crotalariae;			
OC	Crotalaria.			
OX	NCBI_TaxID=3830;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Seed;			
RX	MEDLINE=90167102; PubMed=2306467;			
RA	Khang N.O., Guillaume J.-L., Hoebeke J.;			
RT	"A blood group A specific lectin from the seeds of Crotalaria			
RT	striata.";			
RL	Biochim. Biophys. Acta 1033:210-213(1990).			
CC	-1- FUNCTION: AGGLUTINATES ERYTHROCYTES OF BLOOD GROUP A. BINDS IN			
CC	DECREASING ORDER OF AFFINITY: N-ACETYL-D-GALACTOSAMINE,			
CC	D-GALACTOSE, AND D-GALACTOSAMINE.			
CC	-1- SUBUNIT: Homotetramer.			
CC	-1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.			
DR	PIR: S08292; S08292.			
DR	Interpro: IPR000985; Lectin_lega.			
DR	Interpro: IPR001220; Lectin_legb.			
DR	Pfam: PF00138; Lectin_legb.1.			
DR	PROSITE: PS00307; LECTIN_LEGUMI_BETA; PARTIAL.			
DR	PROSITE: PS00308; LECTIN_LEGUMI_ALPHA; PARTIAL.			
KW	Lectin; Glycoprotein.			
FT	NON_TER			
SQ	SEQUENCE 24 AA; 2875 MW; DC62B82CD9F9BB66 CRC64;			
Query Match	100.0%;	Score 20;	DB 1;	Length 24;
Best Local Similarity	100.0%;	Pred. No. 38;		
Matches	4; Conservative 0;	Mismatches	0;	Indels 0; Gaps 0;
Yy	1 SF5F 4			
Db	5 SF5F 8			
RESULT 2	LEC_DOLAX	STANDARD;	PRT;	32 AA.
ID	P02875;			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Lectin (Fragment).			
DE	Dolichos axillaris (Macrotyloma axillare).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 OC Macrotyloma.
 OC NCBI_TaxID=3876;
 RN [1]
 RP SEQUENCE.
 RA Haylett T., Swart L.S.;
 RT "Isolation and characterization of an anti-A1 lectin from Macrotyloma
 axillare";
 RL S. Afr. J. Chem. 35:33-36(1982).
 CC -1- FUNCTION: METALLOPROTEIN, CONTAINING CA, MG, MN, AND ZN AND
 THE CARBOHYDRATES GALACTOSE, GLUCOSAMINE, MANNOSE, AND FUCCOSE. IT
 AGGLUTINATES ERYTHROCYTES OF BLOOD GROUP A1.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SIMILARITY: BELONGS TO THE LEGUMINOS LECTIN FAMILY.
 DR PIR: A03367; A03367.
 DR HSSP: P05045; 1U1.
 DR InterPro: IPR000985; Lectin_lega.
 DR InterPro: IPR01220; Lectin_legb.
 DR Pfam: PF00139; Lectin_legb; 1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; PARTIAL.
 DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; PARTIAL.
 KW Lectin; Glycoprotein; Calcium.
 FT NON TER 32
 SQ SEQUENCE 32 AA; 3444 MW; D768CE45IDE22A7D CRC64;

Query Match 100.0%; Score 20; DB 1; Length 32;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 5 SFSF 8

RESULT 3

PSAJ_NEPOL STANDARD; PRT; 42 AA.
 AC Q9TKZ8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Photosystem I reaction center subunit IX (PSI-J).
 GN PSAJ.
 OS Nephroselmis olivacea.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
 OC Chlorodendrales; Chlorodendraceae; Nephroselmis.
 OX NCBI_TaxID=31312;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIES-484;
 RX MEDLINE-99398694; PubMed-10468594;
 RA Turnel M., Ots C., Lemieux C.;
 RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
 olivacea: Insights into the architecture of ancestral chloroplast
 genomes";
 RT Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
 CC -1- FUNCTION: May help in the organization of the psae and psaf
 subunits (by similarity).
 CC -1- SIMILARITY: Belongs to the psaf family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AF137379; AAD54818.1; -
 DR HAMAP: MF_00522; -; 1.
 DR InterPro: IPR002615; PSJ_Psaj.

DR Pfam: PF01701; PSI_Psaj; 1.
 DR ProDom: PD004198; PSI_Psaj; 1.
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
 FT TRANSMEM 7
 FT POTENTIAL.
 SQ SEQUENCE 42 AA; 4717 MW; DB684CE1AE903494 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 42;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 38 SFSF 41

RESULT 4

VG66_BPPH2 STANDARD; PRT; 54 AA.
 ID VG66_BPPH2
 AC P16516;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DE Early protein GP16.6.
 GN 16.6.
 OS Bacteriophage phi-29.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC phi-29-like viruses.
 OX NCBI_TaxID=10756;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86165872; PubMed-3007295;
 RA Garvey K.J., Yoshikawa H., Ito J.;
 RT "The complete sequence of the Bacillus phage phi 29 right early
 gene 40:301-309(1985)."
 RL Gene 40:301-309(1985).
 CC -1- SIMILARITY: HIGH, TO EARLY PROTEIN GP16.6 OF BACTERIOPHAGES
 PHI-15 AND PZA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M14430; AAA88353.1; -
 DR PIR: JN0034; JN0034.
 KW Early protein.
 SQ SEQUENCE 54 AA; 6199 MW; 17118C6AC436673B CRC64;

Query Match 100.0%; Score 20; DB 1; Length 54;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 ||||
 Db 12 SFSF 15

RESULT 5

VG66_BPPH5 STANDARD; PRT; 54 AA.
 ID VG66_BPPH5
 AC P15854;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DE Early protein GP16.6.
 GN 16.6.
 OS Bacteriophage phi-15.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC phi-29-like viruses.
 OX NCBI_TaxID=10755;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69232766; PubMed=2497055;
RA Benes V., Arnold L., Smrt J., Paces V.;
RT "Nucleotide sequence of the right early region of Bacillus phage phi
RT 15 and comparison with related phages: reorganization of gene 17
RT during evolution.";
RL Gene 75:341-347(1989).
CC -1 SIMILARITY: HIGH, TO EARLY PROTEIN GP16.6 OF BACTERIOPHAGES
CC PH1-15 AND P2A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1813; AAA88495.1; -
CC DR PIR; B29004; WRBP6.
CC KW Early protein.
SQ SEQUENCE 54 AA; 6114 MW; 57703F8B02E066FF CRC64;

Query Match 100.0%; Score 20; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
VG66_BPPZA
ID VG66_BPPZA STANDARD; PRT; 54 AA.
AC P08385;
DT 01-AUG-1988 (rel. 08. Created)
DT 01-AUG-1988 (rel. 08. Last sequence update)
DT 01-APR-1990 (rel. 14. Last annotation update)
DE Early protein GP16.6.
GN 16.6.
OS Bacteriophage P2A.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OX NCB1_TaxID=10757;
RX NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=87031575; PubMed=3095189;
RA Paces V., Vlcek C., Urbnek P., Hostomsky Z.;
RT "Nucleotide sequence of the right early region of Bacillus subtilis
RT phage P2A completes the 19366-bp sequence of P2A genome. Comparison
RT with the homologous sequence of phage phi 29.";
RL Gene 44:115-120(1986).
CC -1 SIMILARITY: HIGH, TO EARLY PROTEIN GP16.6 OF BACTERIOPHAGES
CC PH1-29 AND PH1-15.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1813; AAA88495.1; -
CC DR PIR; B29004; WRBP6.
CC KW Early protein.
SQ SEQUENCE 54 AA; 6114 MW; 57703F8B02E066FF CRC64;

Query Match 100.0%; Score 20; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1 SF5F 4
        ||||
Db      12 SF5F 15

RESULT 7
YIMB_BPPH1
ID      YIMB_BPPH1      STANDARD;      PRT;      57 AA.
AC      P10434;
DT      01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      01-MAR-1989 (Rel. 10, Last annotation update)
DE      Hypothetical immunity region protein 11.
OS      Bacteriophage phi-105.
OC      Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC      Lambda-like viruses.
OX      NCBI_TaxID=10717;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86056972; PubMed=3934047;
RA      Cully D.F., Garro A.J.;
RT      "Nucleotide sequence of the immunity region of Bacillus subtilis
RT      bacteriophage phi 105: Identification of the repressor gene and its
RT      mRNA and protein products."
RL      Gene 38:153-164(1985).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL, M11920; AAA88398.1; -.
DR      EMBL; A27234; IMBP11.
DR      PIR; A27234; IMBP11.
KW      Hypothetical protein.
SQ      SEQUENCE 57 AA; 6636 MW; 9EC83261EB3D4EE6 CRC64;

Query Match      100.0%; Score 20; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Oy      1 SF5F 4
        ||||
Db      6 SF5F 9

RESULT 8
BP4A_BRANA
ID      BP4A_BRANA      STANDARD;      PRT;      63 AA.
AC      P41505;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      01-NOV-1995 (Rel. 32, Last annotation update)
DE      BP4A protein.
GN      BP4A.
OS      Brassica napus (Rape).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Brassica.
OX      NCBI_TaxID=3708;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=cv. Westar.
RX      MEDLINE=91338699; PubMed=2102378;
RA      Albari D., Robert L.S., Donaldson P.A., Altsaer I., Arnison P.G.,
RA      Fabijanski S.F.;
RT      "Characterization of a pollen-specific gene family from Brassica
RT      napus which is activated during early microspore development.";
RT      Plant Mol. Biol. 15:605-622(1990).
CC      -1-TISSUE SPECIFICITY: Pollen-specific.

```

CC -1- DEVELOPMENTAL STAGE: ACTIVATED DURING EARLY MICROSPORE
CC DEVELOPMENT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52874; CAA37053.1; -
DR PIR; S12241; S12241.
KW Multigene family.

SO SEQUENCE 63 AA; 6883 MW; CAD55EAB798C44B1 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 8 SFSF 11

RESULT 9

YVAP_VACC STANDARD; PRT; 89 AA.

AC P20525;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 9.9 kDa protein.

GN A ORF P.

OS Vaccinia virus (strain Copenhagen).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus.

OX NCBI_TaxID=10249;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91021027; PubMed=2219722;

RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,

RA Paolletti E.;

RT "The complete DNA sequence of vaccinia virus.";

RL Virology 179:247-266(1990).

CC COMPLETE GENOME.

RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,

RA Paolletti E.;

RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";

RL Virology 179:517-563(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M35027; AAA48170.1; -

DR PIR; C42525; C42525.

KW Hypothetical protein.

SO SEQUENCE 89 AA; 9909 MW; 399EA9270DEFE73A CRC64;

Query Match 100.0%; Score 20; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 10 SFSF 13

RESULT 10

YCXF_CYAPA STANDARD; PRT; 91 AA.

AC P48336;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 11.2 kDa protein in YCF23-APCF intergenic region (ORF91).

OS Cyanophora paradoxa.

OC Cyanelle.

OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.

OX NCBI_TaxID=2762;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-UTEX LB 555 / Pringsheim;

RA Loeffelhardt W., Stirewalt V.L., Loeffelhardt W., Bohnert H.J.,

RA Stirewalt V.L., Michalowski C.B., Bryant D.A.;

RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";

RL Plant Mol. Biol. Rep. 13:327-332(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-UTEX LB 555 / Pringsheim;

RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,

RA Farley J.V., Schluchter W.M., Chung S., Newmann-Spallart C.,

RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;

RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:

the genetic complexity of a primitive plastid.";

RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,

RL Schwemmer W. (eds.);

RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg

CC (1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U30821; AAA81306.1; -

DR PIR; T06963; T06963.

KW Hypothetical protein; Cyanelle.

SO SEQUENCE 91 AA; 11166 MW; 61E65208C9DD4D6F CRC64;

Query Match 100.0%; Score 20; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSF 4
1111
DB 19 SFSF 22

RESULT 11
POL_HVLJ3 STANDARD; PRT; 100 AA.

AC P12498;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE POL polyprotein (Contains: Protease (Retropepsin) (EC 3.4.23.16);

DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4))

DE (Fragment).

OS POL.

OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11694;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89352108; PubMed=2669897;

RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;

```

RT      "Nucleotide sequences of gag and env genes of a Japanese isolate of
RT      HIV-1 and their expression in bacteria."
RT      AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC      -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC      and p1' variable, but often Pro.
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC      phosphonomonoester.
CC      -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC      + (dNp)(N).
CC      -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
CC      DETERMINED.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC      KNOWN AS THE RETROPEPSIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M21137; AAB03523.1; -.
CC      DR      HIV; M21137; POLS3H3.
CC      DR      MEROPS; A02.001; -.
CC      DR      InterPro: IPR001969; Aspprotease_rtrv.
CC      DR      InterPro: IPR001969; Aspprotease_site.
CC      DR      Pfam; PF00077; rvp; 1.
CC      DR      PROSITE; PS00141; ASP_PROTEASE; 1.
CC      DR      PROSITE; PS00175; ASP_PROT_RETROV; 1.
CC      DR      AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC      KW      Nuclease; Transferase; RNA-directed DNA polymerase.
CC      FT      CHAIN 57 >100
CC      FT      ACT_SITE 81 81 BY SIMILARITY.
CC      FT      NON_TER 100 100
CC      SQ      SEQUENCE 100 AA; 11208 MW; C90A958EDE327254 CRC64;
CC
CC      Query Match          100.0%; Score 20; DB 1; Length 100;
CC      Best Local Similarity 100.0%; Pred. No. 1.6e+02;
CC      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RT      Science 273:1058-1073(1996).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U67464; AAB98052.1; -.
CC      DR      PIR; C64308; C64308.
CC      DR      TIGR; MJ0067; -.
CC      DR      InterPro: IPR007014; FUN14.
CC      DR      Pfam; PF04930; FUN14; 1.
CC      DR      Hypothetical protein; Transmembrane; Complete proteome.
CC      KW      TRANSMEM 6 26 POTENTIAL.
CC      FT      TRANSMEM 34 54 POTENTIAL.
CC      FT      TRANSMEM 79 99 POTENTIAL.
CC      SQ      SEQUENCE 103 AA; 11059 MW; 94DCD5C2P394D53 CRC64;
CC
CC      Query Match          100.0%; Score 20; DB 1; Length 103;
CC      Best Local Similarity 100.0%; Pred. No. 1.6e+02;
CC      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

SQ SEQUENCE 110 AA; 13187 MW; B905F6BFBFF207DF CRC64;

Query Match

Best Local Similarity 100.0%; Score 20; DB 1; Length 110;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSE 4
1111
Db 40 SFSE 43

RESULT 14

YC4_YEAST STANDARD; PRT; 111 AA.

AC P53190;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 12.2 kDa protein in PGD1-STT3 intergenic region.
GN YGL024W.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

RN [1]
RP SEQUENCE FROM N.A.

RA Hebling U., Hofmann B., Delius H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 272547; CAA96726.1; -.
DR PIR; S64026; S64026.
DR SGD; S0002992; YGL024W.

KW Hypothetical protein; Transmembrane.

FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.

SQ SEQUENCE 111 AA; 12210 MW; D31EAB5717F09C22 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSE 4
1111
Db 34 SFSE 37

RESULT 15

Y465_AQUAE STANDARD; PRT; 113 AA.

AC 066767;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ_465.

GN AQ_465.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN-VF5;

RA MEDLINE=98196666; PubMed=9537320;
Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anujay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus.";

RT Nature 392:353-358(1998).

RL Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE000691; AAC06729.1; -.

DR PIR; C70342; C70342.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 7 29 POTENTIAL.

SQ SEQUENCE 113 AA; 13252 MW; F402B9193B2A8522 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSE 4
1111
Db 73 SFSE 76

Search completed: August 20, 2003, 12:34:56
Job time : 4.62918 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 7.56627 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-30
Perfect score: 20
Sequence: 1 SFSF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SPTREMBL_23:*

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	10 Q9S8J5	Q9S8J5 psophocarpus
2	20	100.0	20	10 Q9S8B2	Q9S8B2 phaseolus v
3	20	100.0	23	4 Q9G12	Q9G12 homo sapien
4	20	100.0	31	2 Q4547	Q4547 bacillus su
5	20	100.0	36	4 Q8N606	Q8N606 homo sapien
6	20	100.0	40	10 Q9S8X8	Q9S8X8 erythrina v
7	20	100.0	41	12 Q91FB5	Q91FB5 chilo iride
8	20	100.0	42	2 Q49491	Q49491 mycoplasma
9	20	100.0	42	16 Q9K8D7	Q9K8D7 bacillus ha
10	20	100.0	44	16 Q8E2R7	Q8E2R7 leptospira
11	20	100.0	46	8 Q8SM41	Q8SM41 spatoglossu
12	20	100.0	48	3 Q60031	Q60031 aspergillus
13	20	100.0	48	16 Q8K18	Q8K18 chlorobium
14	20	100.0	49	2 Q9FB83	Q9FB83 helicobacte
15	20	100.0	49	2 Q9FB93	Q9FB93 helicobacte
16	20	100.0	50	16 Q9K1LO	Q9K1LO vibrio chol

17	20	100.0	50	16 Q8U5A9	Q8U5A9 agrobacteri
18	20	100.0	51	2 Q05977	Q05977 synechocyst
19	20	100.0	51	16 Q9K0R3	Q9K0R3 nisseria m
20	20	100.0	52	2 Q57376	Q57376 leptospira
21	20	100.0	52	16 Q8E2Z7	Q8E2Z7 leptospira
22	20	100.0	53	2 Q9X559	Q9X559 enterococu
23	20	100.0	55	8 P92500	P92500 apis mellif
24	20	100.0	55	8 P92499	P92499 apis mellif
25	20	100.0	55	8 P92501	P92501 apis mellif
26	20	100.0	55	8 Q8H003	Q8H003 thrips imag
27	20	100.0	55	10 Q8G202	Q8G202 arabidopsis
28	20	100.0	57	2 Q48855	Q48855 lactobacill
29	20	100.0	57	12 Q91N00	Q91N00 lumpy skin
30	20	100.0	58	5 Q24855	Q24855 entamoeba h
31	20	100.0	59	5 Q8MS44	Q8MS44 drosophila
32	20	100.0	60	5 Q8T2D6	Q8T2D6 dictyostell
33	20	100.0	61	10 Q92VBA	Q92VBA arabidopsis
34	20	100.0	61	16 Q8E9E8	Q8E9E8 oceanobact
35	20	100.0	62	5 Q81IF3	Q81IF3 plasmodium
36	20	100.0	62	10 Q9M7N2	Q9M7N2 picroamia p
37	20	100.0	62	16 Q8F7T7	Q8F7T7 leptospira
38	20	100.0	63	15 Q8AEV7	Q8AEV7 human immun
39	20	100.0	63	16 Q9PB3	Q9PB3 xylella fas
40	20	100.0	64	15 Q8AEV9	Q8AEV9 human immun
41	20	100.0	64	15 Q8AEV5	Q8AEV5 human immun
42	20	100.0	64	16 Q9J223	Q9J223 nisseria m
43	20	100.0	64	17 Q8PTW2	Q8PTW2 methanosarc
44	20	100.0	65	5 Q8IEZ7	Q8IEZ7 trypanosoma
45	20	100.0	65	11 Q8BW43	Q8BW43 mus musculu

ALIGNMENTS

RESULT 1

AC Q9S8J5 PRELIMINARY; PRT: 20 AA.
ID Q9S8J5
DR 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Lectin 1 (Fragment).
OS Psophocarpus tetragonolobus (Goa bean) (Asparagus bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OX NCBI_TaxID=3891;
RN [1]
RP MEDLINE=95120285; PubMed=7820375;
RA Yagi F., Sawada R., Imada T., Toyonaga S., Tadara K., Ishihata K.;
RT "Two isolectins from leaves of winged bean, Psophocarpus
tetragonolobus (L.) DC.";
RL Plant Cell Physiol. 35:1087-1095(1994).
SQ SEQUENCE 20 AA; 2363 MW; 1BF1B347020D3DCE CRC64;

Query Match

Best Local Similarity 100.0%; Score 20; DB 10; Length 20;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy. 1 SFSF 4
1111
DB 5 SFSF 8

RESULT 2

AC Q9S8B2 PRELIMINARY; PRT: 20 AA.
ID Q9S8B2
DR 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE GNL-2 alpha subunit (Fragment).
 OS Phaseolus vulgaris (Kidney bean). (French bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxID=3885;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96203526; PubMed=8605238;
 RA Kamemura K., Furuchi Y., Umekawa H., Takahashi T.;
 RT "Purification and characterization of a pod lectin from Great Northern
 bean, Phaseolus vulgaris L.";
 RL Biochim. Biophys. Acta 1289:87-94(1996).
 SQ SEQUENCE 20 AA; 2399 MW; B09E83F93CCD833 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 SFSF 4
 Db 5 SFSF 8

RESULT 3
 096G12 PRELIMINARY; PRT; 23 AA.
 AC 096G12;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC010062; AAH10062.1; -
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 23 AA; 2476 MW; 36E4FA8D2C7289A CRC64;

Query Match 100.0%; Score 20; DB 4; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFSF 4
 Db 5 SFSF 8

RESULT 4
 045547 PRELIMINARY; PRT; 31 AA.
 AC 045547;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Divin (Fragment).
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M23;
 RA Harry E.J., Partridge S.R., Wake R.G.;
 RT "divin region of Bacillus licheniformis.";
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L25649; AAA2410.1; -

FT NON_TER 31
 SQ SEQUENCE 31 AA; 3434 MW; 01B2433555A76D31 CRC64;

Query Match 100.0%; Score 20; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFSF 4
 Db 21 SFSF 24

RESULT 5
 08N606 PRELIMINARY; PRT; 36 AA.
 AC 08N606;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Similar to hypothetical protein BC010062 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC029129; AAH29129.1; -
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 36 AA; 3704 MW; 6935F10917DDE61 CRC64;

Query Match 100.0%; Score 20; DB 4; Length 36;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFSF 4
 Db 18 SFSF 21

RESULT 6
 09S8X8 PRELIMINARY; PRT; 40 AA.
 AC 09S8X8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
 DE Agglutinin (Fragment).
 OS Erythrina variegata (Coral tree).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Erythrina.
 OX NCBI_TaxID=3845;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92361170; PubMed=2136348;
 RA Li H., Yamamoto K., Kawashima H., Osawa T.;
 RT "Structural requirements for the binding of oligosaccharides to
 RT immobilized lectin of Erythrina variegata (Lin) var. orientalis.";
 RL Glycoconj. J. 7:311-322(1990).
 DR HSSP; P16404; IAX2.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00139; Lectin_legB; 1.
 DR Pfam; PD000711; Lectin_legB; 1.
 SQ SEQUENCE 40 AA; 4320 MW; 0CD3D402FAC5B556 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 40;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFSF 4
 ||||
 Db 5 SFSF 8

RESULT 7
 Q91FB5 PRELIMINARY; PRT; 41 AA.
 AC Q91FB5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE 409R.
 OS Chilo Iridescent virus (CIV) (Insect Iridescent virus type 6).
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
 OX NCBI_TaxID=10488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dellus H., Darai G., Flugel R.M.;
 RT "DNA analysis of insect Iridescent virus 6: evidence for circular
 permutational and terminal redundancy.";
 RL J. Virol. 49:609-614(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86174607; PubMed=3959991;
 RA Iorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
 RT "Insect Iridescent virus type 6 induced toxic degenerative hepatitis
 in mice.";
 RL Med. Microbiol. Immunol. 175:43-53(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87321126; PubMed=2820141;
 RA Schiltzler P., Soltau J.B., Fischer M., Reisner H., Scholz J.,
 RA Dellus H., Darai G.;
 RT "Molecular cloning and physical mapping of the genome of insect
 Iridescent virus type 6: further evidence for circular permutation of
 the viral genome.";
 RL Virology 160:66-74(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89073752; PubMed=3201750;
 RA Fischer M., Schiltzler P., Dellus H., Darai G.;
 RT "Identification and characterization of the repetitive DNA element in
 the genome of insect Iridescent virus type 6.";
 RL Virology 167:485-496(1988).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92196996; PubMed=1549908;
 RA Handermann M., Schiltzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 RA Darai G.;
 RT "Identification and mapping of origins of DNA replication within the
 DNA sequences of the genome of insect Iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93118242; PubMed=1475907;
 RA Sonntag K.C., Darai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 of insect Iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93260401; PubMed=8492091;
 RA Stohwasser R., Raab K., Schiltzler P., Janssen W., Darai G.;
 RT "Identification of the gene encoding the major capsid protein of
 insect Iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94167241; PubMed=8121799;
 RA Schiltzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
 RA Dellus H., Darai G.;
 RT "Identification of genes encoding zinc finger proteins, non-histone

RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
 RT in the genome of Chilo Iridescent virus.";
 RL Nucleic Acids Res. 22:158-166(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94353641; PubMed=8073636;
 RA Sonntag K.C., Schiltzler P., Koonin E.V., Darai G.;
 RT "Chilo Iridescent virus encodes a putative helicase belonging to a
 distinct family within the 'Dead/H' superfamily: implications for the
 evolution of large DNA viruses.";
 RL Virus Genes 8:151-158(1994).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95213160; PubMed=7698884;
 RA Sonntag K.C., Schiltzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 the genome of insect Iridescent virus type 6 between the genome
 coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292906; PubMed=8021587;
 RA Schiltzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 RA Koonin E.V., Darai G.;
 RT "Insect Iridescent virus type 6 encodes a polypeptide related to the
 largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo Iridescent virus between the genome
 coordinates 0.101 and 0.391: similarities in coding strategy between
 insect and vertebrate Iridoviruses.";
 RL Virus Genes 15:235-245(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99125223; PubMed=9926400;
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 Chilo Iridescent virus.";
 RL Virus Genes 17:243-258(1998).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99383793; PubMed=10456793;
 RA Muller K., Tidona C.A., Darai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 Iridescent virus encoding enzymes involved in viral DNA replication
 and processing.";
 RL Virus Genes 18:243-264(1999).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21342589; PubMed=11448171;
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT "Analysis of the First Complete DNA Sequence of an Invertebrate
 Iridovirus: Coding Strategy of the genome of Chilo Iridescent Virus.";
 RL Virology 286:182-196(2001).
 RN [16]
 RP SEQUENCE FROM N.A.
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL AF303741; AK82269.1; 30FE367A155F2C8A CRC64;
 SQ SEQUENCE 41 AA; 4991 MW; 30FE367A155F2C8A CRC64;

Query Match 100.0%; Score 20; DB 12; Length 41;
 Best local similarity 100.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFSF 4
 ||||
 Db 7 SFSF 10

RESULT 8

Q49491 PRELIMINARY; PRT; 42 AA.
 ID 049491
 AC 049491
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Random genomic sequence MG44 (Fragment).
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 NX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=92051396; PubMed=1945886;
 RA MEDLINE=92051396; PubMed=1945886;
 RA Peterson S.N., Schramm N., Hu P.C., Bott R.F., Hutchison C.A.;
 RT "A random sequencing approach for placing markers on the physical map
 of Mycoplasma genitalium.";
 RL Nucleic Acids Res. 19:6027-6031(1991).
 DR EMBL; X61539; CAA43751.1; -.
 FT NON_TER 1
 FT NON_TER 42
 SQ SEQUENCE 42 AA; 5110 MW; 07BFC4EAF9C577F CRC64;

Query Match 100.0%; Score 20; DB 2; Length 42;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 26 SFSF 29

RESULT 9

Q9K8D7 PRELIMINARY; PRT; 42 AA.
 ID 09K8D7
 AC 09K8D7
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein BH3069.
 GN BH3069.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RA MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001517; BAB06788.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 42 AA; 5136 MW; 64021B12609B5703 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 42;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 35 SFSF 38

RESULT 10

Q8EZ7 PRELIMINARY; PRT; 44 AA.
 ID 08EZ7
 AC 08EZ7
 DT 01-MAR-2003 (TREMblrel. 23, Created)

DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN LA3785.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 NX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011534; AAN50983.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 44 AA; 5239 MW; B3374E64BE8A1D6 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 44;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 30 SFSF 33

RESULT 11

Q8SM41 PRELIMINARY; PRT; 46 AA.
 ID 08SM41
 AC 08SM41
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Ribulose-1,5-bisphosphate carboxylase small subunit (Fragment).
 GN RBCS.
 OS Spatoglossum crassum.
 OC Chloroplast.
 OC Eukaryota; stramenopiles; Phaeophyceae; Dictyotales; Dictyotaceae;
 OC Spatoglossum.
 NX NCBI_TaxID=157007;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee W.J., Bae K.S.;
 RT "A phylogenetic relationship among the genera of Dictyotaceae
 (Dictyotales, Phaeophyta) based on partial rbcL sequences.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF353382; AAM08198.1; -.
 DR InterPro: IPR002160; Kunitz_legume.
 DR InterPro: IPR000894; Rubisco_small.
 DR Pfam: PF00101; Rubisco_small; 1.
 DR ProDom: PD000290; Rubisco_small; 1.
 DR PROSITE: PS00283; SOYBEAN_KUNITZ; 1.
 KW Chloroplast.
 FT NON_TER 46
 SQ SEQUENCE 46 AA; 5339 MW; 0B6B2F28BBA931CF CRC64;

Query Match 100.0%; Score 20; DB 8; Length 46;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 7 SFSF 10

RESULT 12

Q60031 PRELIMINARY; PRT; 48 AA.
 ID 060031
 AC 060031
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical 5.7 kda protein.
 OS Aspergillus niger.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95374684; PubMed=7646819;
 RA van den Brink J., van Zeeijl C., Brons J., van den Hondel C.,
 van Gorcom R.;
 RT "Cloning and characterization of the NADPH-cytochrome P450
 RT oxidoreductase gene from the filamentous fungus *Aspergillus niger*.";
 RL DNA Cell Biol. 14:719-729(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20309321; PubMed=10852481;
 RA van den Brink J.M., Punt P.J., Van Gorcom R.F.M.,
 van den Hondel C.A.M.;
 RT "Regulation of expression of the *Aspergillus niger* benzoate para-
 RT hydroxylase cytochrome P450 system.";
 RL Mol. Gen. Genet. 263:601-609(2000).
 DR EMBL: AJ005117; CA06381.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 48 AA; 5746 MW; EB85F2B8CA0B1CDD CRC64;

Query Match 100.0%; Score 20; DB 3; Length 48;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 24 SFSF 27

RESULT 13

Q8KB18 PRELIMINARY; PRT; 48 AA.
 AC Q8KB18;
 DT 01-OCT-2002 (TREMblrel. 22, Created)
 DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Hypothetical protein CT1980.
 GN CT1980.
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 CC Chlorobium.
 OX NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-TLS / ATCC 49652 / DSM 12025;
 RC MEDLINE=22103685; PubMed=12093901;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
 Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
 Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
 Nierman W.C., Feldblyum T.V., Hansen C.L., Craen M.B., Radune D.,
 Vamathevan J., Khouli H., White O., Gruber T.M., Ketchum K.A.,
 Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
 RT "The complete genome sequence of *Chlorobium tepidum* TLS, a
 RT photosynthetic, anaerobic, green-sulfur bacterium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL: AE012948; AAW3198.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 48 AA; 5801 MW; 4A3D5E7BE1A477DD CRC64;

Query Match 100.0%; Score 20; DB 16; Length 48;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 35 SFSF 38

RESULT 14

Q9FB83 PRELIMINARY; PRT; 49 AA.
 AC Q9FB83;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN OMPA.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 CC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-JK2-41;
 RC MEDLINE=20319061; PubMed=10852959;
 RA Yamaoka Y., Kwon D.H., Graham D.Y.;
 RT "A Mrp 34,000 proinflammatory outer membrane protein (ompA) of
 RT Helicobacter pylori.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7533-7538(2000).
 DR EMBL: AF233670; AAG00390.1; -;
 KW NON_TER
 FT
 SQ SEQUENCE 49 AA; 5372 MW; FEE2E56DD1310AC0 CRC64;

Query Match 100.0%; Score 20; DB 2; Length 49;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 10 SFSF 13

RESULT 15

Q9FB93 PRELIMINARY; PRT; 49 AA.
 AC Q9FB93;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN OMPA.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 CC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-JK23;
 RC MEDLINE=20319061; PubMed=10852959;
 RA Yamaoka Y., Kwon D.H., Graham D.Y.;
 RT "A Mrp 34,000 proinflammatory outer membrane protein (ompA) of
 RT Helicobacter pylori.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7533-7538(2000).
 DR EMBL: AF233660; AAG00380.1; -;
 KW NON_TER
 FT
 SQ SEQUENCE 49 AA; 5471 MW; 43442D9DD131019E CRC64;

Query Match 100.0%; Score 20; DB 2; Length 49;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSF 4
 DB 10 SFSF 13

Search completed: August 20, 2003, 12:40:27
 Job time : 12.6774 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 ; Search time 8.79518 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26
Sequence: 1 PPFY 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	9	21	AA10067
2	26	100.0	10	21	AA10068
3	26	100.0	11	21	AA10069
4	26	100.0	12	21	AA10070
5	26	100.0	13	21	AA10071
6	26	100.0	13	21	AA10072
7	26	100.0	13	22	AA10073
8	26	100.0	15	22	AA10074
9	26	100.0	15	23	AA10075

10	26	100.0	16	23	ABP46251
11	26	100.0	20	19	AAW48100
12	26	100.0	22	22	AA677700
13	26	100.0	23	22	AA666072
14	26	100.0	27	18	AAW19473
15	26	100.0	27	18	AAW19478
16	26	100.0	28	15	AAW59870
17	26	100.0	32	23	ABP30850
18	26	100.0	36	21	AA659652
19	26	100.0	36	22	AAU03617
20	26	100.0	38	17	AAW89894
21	26	100.0	38	18	AAW19427
22	26	100.0	38	18	AAW19444
23	26	100.0	38	19	AAW31678
24	26	100.0	39	15	AAW58363
25	26	100.0	39	22	ABG59752
26	26	100.0	39	22	AAW65444
27	26	100.0	39	22	AAW38348
28	26	100.0	47	21	AAW91367
29	26	100.0	48	21	AAW24535
30	26	100.0	50	22	AAO10716
31	26	100.0	51	22	AAW14820
32	26	100.0	51	22	AAW97974
33	26	100.0	51	22	AAW42609
34	26	100.0	51	23	ABP35458
35	26	100.0	52	21	AAW55535
36	26	100.0	53	19	AAW74848
37	26	100.0	53	23	ABG95299
38	26	100.0	56	20	AAW48523
39	26	100.0	59	22	AAU20693
40	26	100.0	60	24	ABP76248
41	26	100.0	62	24	ABP75387
42	26	100.0	66	23	ABP39857
43	26	100.0	66	23	ABP08962
44	26	100.0	68	22	AAW92354
45	26	100.0	68	22	AAU20050

ALIGNMENTS

RESULT 1	AA10067	standard; peptide; 9 AA.
ID	AA10067	
XX	AA10067	
AC	AA10067	
DT	03-NOV-2000	(first entry)
XX		
DE	U1A protein derived tryptic peptide #6.	
XX		
KW	U1A protein; tryptic peptide; functional site; detection.	
XX		
OS	Unidentified.	
XX		
PN	WO200031544-A1.	
XX		
PD	02-JUN-2000.	
XX		
PF	23-NOV-1999; 99WO-EP09052.	
XX		
PR	24-NOV-1998; 98DE-1054196.	
XX		
PA	(XERI-) XERION PHARM GMBH.	
XX		
PI	Ilag IL, Ng JH;	
XX		
DR	WPI; 2000-400193/34.	
XX		
PT	Identifying functional sites in proteins, useful for detecting epitopes	
PT	and ligand binding sites, by complexing with tagged binding partner,	
PT	then laser irradiation to alter the protein close to the tag	
XX		

PS Example 2; Page 24; 41pp; German.
 CC This invention describes a novel method for identifying one or more
 CC functional sites in a target protein (I) which comprises: (i) complexing
 CC (I) with a binding partner (A) that carries a laser-activatable tag; (ii)
 CC irradiating the complex (II) formed with laser light to generate free
 CC radicals that alter bound (I) selectively at the binding sites; and (iii)
 CC identifying the altered region by a combination of protein cleavage and
 CC mass spectrometry (MS). The method is used to identify any type of
 CC functional site in any type of protein but particularly ligand-binding
 CC sites and epitopes. It can also be used to investigate pathologically
 CC altered proteins or oncogenic proteins, and to inactivate target
 CC pathological proteins. The method does not require knowledge of the
 CC three-dimensional structure of (I), and is simple, rapid and automatable.
 CC It can identify non-linear or discontinuous epitopes, and determines
 CC protein function without inactivation of (I) (since the tag is bound to
 CC (A) not to (I) itself). The method may be used in vitro or in vivo for
 CC studying intra- or extra-cellular proteins. This sequence represents a
 CC method of the invention.

SQ Sequence 9 AA;

Query Match 100.0%; Score 26; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 DB 1 FPFY 4

RESULT 2
 AAB10068
 ID AAB10068 standard; peptide; 10 AA.

XX AAB10068;

AC 03-NOV-2000 (first entry)

DE U1A protein derived tryptic peptide #7.

XX U1A protein; tryptic peptide; functional site; detection.

OS Unidentified.

XX WO200031544-A1.

PN 02-JUN-2000.

PD 23-NOV-1999; 99WO-EP09052.

PR 24-NOV-1998; 98DE-1054196.

XX (XERT-) XERTON PHARM GMBH.

PA I1ag LL, Ng JH;

PI WPI; 2000-400193/34.

DR Identifying functional sites in proteins, useful for detecting epitopes
 XX and ligand binding sites, by complexing with tagged binding partner,
 PT then laser irradiation to alter the protein close to the tag
 PS Example 2; Page 24; 41pp; German.

CC This invention describes a novel method for identifying one or more
 CC functional sites in a target protein (I) which comprises: (i) complexing
 CC (I) with a binding partner (A) that carries a laser-activatable tag; (ii)
 CC irradiating the complex (II) formed with laser light to generate free
 CC radicals that alter bound (I) selectively at the binding sites; and (iii)
 CC identifying the altered region by a combination of protein cleavage and
 CC mass spectrometry (MS). The method is used to identify any type of

CC functional site in any type of protein but particularly ligand-binding
 CC sites and epitopes. It can also be used to investigate pathologically
 CC altered proteins or oncogenic proteins, and to inactivate target
 CC pathological proteins. The method does not require knowledge of the
 CC three-dimensional structure of (I), and is simple, rapid and automatable.
 CC It can identify non-linear or discontinuous epitopes, and determines
 CC protein function without inactivation of (I) (since the tag is bound to
 CC (A) not to (I) itself). The method may be used in vitro or in vivo for
 CC studying intra- or extra-cellular proteins. This sequence represents a
 CC method of the invention.

SQ Sequence 10 AA;

Query Match 100.0%; Score 26; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 DB 2 FPFY 5

RESULT 3
 AAB10069
 ID AAB10069 standard; peptide; 11 AA.

XX AAB10069;

AC 03-NOV-2000 (first entry)

DE U1A protein derived tryptic peptide #8.

XX U1A protein; tryptic peptide; functional site; detection.

OS Unidentified.

XX WO200031544-A1.

PN 02-JUN-2000.

PD 23-NOV-1999; 99WO-EP09052.

PR 24-NOV-1998; 98DE-1054196.

XX (XERT-) XERTON PHARM GMBH.

PA I1ag LL, Ng JH;

PI WPI; 2000-400193/34.

DR Identifying functional sites in proteins, useful for detecting epitopes
 XX and ligand binding sites, by complexing with tagged binding partner,
 PT then laser irradiation to alter the protein close to the tag
 PS Example 2; Page 24; 41pp; German.

CC This invention describes a novel method for identifying one or more
 CC functional sites in a target protein (I) which comprises: (i) complexing
 CC (I) with a binding partner (A) that carries a laser-activatable tag; (ii)
 CC irradiating the complex (II) formed with laser light to generate free
 CC radicals that alter bound (I) selectively at the binding sites; and (iii)
 CC identifying the altered region by a combination of protein cleavage and
 CC mass spectrometry (MS). The method is used to identify any type of
 CC functional site in any type of protein but particularly ligand-binding
 CC sites and epitopes. It can also be used to investigate pathologically
 CC altered proteins or oncogenic proteins, and to inactivate target
 CC pathological proteins. The method does not require knowledge of the
 CC three-dimensional structure of (I), and is simple, rapid and automatable.
 CC It can identify non-linear or discontinuous epitopes and determines
 CC protein function without inactivation of (I) (since the tag is bound to
 CC (A) not to (I) itself). The method may be used in vitro or in vivo for
 CC studying intra- or extra-cellular proteins. This sequence represents a

CC U1A protein derived tryptic peptide which is used to illustrate the
 CC method of the invention.

SO Sequence 11 AA;

Query Match 100.0%; Score 26; DB 21; Length 11;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 DB 3 FPFY 6

RESULT 4
 AAB10070
 ID AAB10070 standard; peptide; 12 AA.

AC AAB10070;

DT 03-NOV-2000 (first entry)

DE U1A protein derived tryptic peptide #9.

XX U1A protein; tryptic peptide; functional site; detection.

OS Unidentified.

XX WO200031544-A1.

XX 02-JUN-2000.

XX 23-NOV-1999; 99WO-EP09052.

XX 24-NOV-1998; 98DE-1054196.

XX (XERI-) XERION PHARM GMBH.

XX I1ag LI, Ng JH;

XX WPI; 2000-400193/34.

PT Identifying functional sites in proteins, useful for detecting epitopes
 PT and ligand binding sites, by complexing with tagged binding partner,
 PT then laser irradiation to alter the protein close to the tag

PS Example 2; Page 24; 41pp; German.

CC This invention describes a novel method for identifying one or more
 CC functional sites in a target protein (I) which comprises: (i) complexing
 CC (I) with a binding partner (A) that carries a laser-activatable tag; (ii)
 CC irradiating the complex (ii) formed with laser light to generate free
 CC radicals that alter bound (I) selectively at the binding sites; and (iii)
 CC identifying the altered region by a combination of protein cleavage and
 CC mass spectrometry (MS). The method is used to identify any type of
 CC functional site in any type of protein but particularly ligand-binding
 CC sites and epitopes. It can also be used to investigate pathologically
 CC altered proteins or oncogenic proteins, and to inactivate target
 CC pathological proteins. The method does not require knowledge of the
 CC three-dimensional structure of (I), and is simple, rapid and automatable.
 CC It can identify non-linear or discontinuous epitopes, and determines
 CC protein function without inactivation of (I) (since the tag is bound to
 CC (A) not to (I) itself). The method may be used in vitro or in vivo for
 CC studying intra- or extra-cellular proteins. This sequence represents a
 CC U1A protein derived tryptic peptide which is used to illustrate the
 CC method of the invention.

SO Sequence 12 AA;

Query Match 100.0%; Score 26; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 DB 4 FPFY 7

RESULT 5
 AAB10003
 ID AAB10003 standard; protein; 13 AA.

AC AAB10003;

DT 01-NOV-2000 (first entry)

DE H. pylori 26 kDa protein-binding antibody heavy chain CDR3 peptide.

XX Acid-resistant microorganism; detection; faecal; intestine; infection;
 XX monoclonal antibody; heavy chain; complementarity determining region;
 XX CDR.

XX Unidentified.

XX WO200026671-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-EP08212.

XX 29-OCT-1998; 98EP-0120517.

XX 06-NOV-1998; 98EP-0120687.

XX (CONN-) CONNEX GMBH.

XX Relfer C, Cullmann G, Friedrichs U, Heppner P, Lakner M;
 XX Ringels A;

XX WPI; 2000-365747/31.

XX N-PSDB; AAB10003.

PT Detecting infection by acid-fast microbes for diagnosis of Helicobacter
 PT pylori, comprises reacting a faecal sample with two binding reagents for
 PT antigens that survive intestinal passage

PS Claim 22; Page 21; 84pp; German.

CC This invention describes a novel method for the detection of a mammalian
 CC infection by an acid-resistant microorganism (A) by treating a faecal
 CC sample with at least two different monoclonal antibodies (MAb) (or their
 CC fragments or derivatives) or aptamers (collectively (II)) and detecting
 CC formation of a complex (C) between (I) and the corresponding antigen of
 CC (A). The first and second (I) bind to epitopes of different antigens,
 CC (Ag). These epitopes are present, after passage through the intestines,
 CC in at least some mammals, and have either: (i) their native structure;
 CC or (ii) a structure against which an antibody is produced by an animal
 CC infected or immunized with (A), or its extract, lysate, derived protein
 CC or fragment, or with a synthetic peptide. Practically all mammals display
 CC at least one of the specified epitopes. The method is used to detect
 CC infection by acid-fast bacteria, particularly of the genera Helicobacter,
 CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,
 CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used
 CC therapeutically. The method is direct and non-invasive, and provides an
 CC inexpensive and easily standardizable diagnosis, despite possible
 CC degradation of antigens during passage through the intestines. This
 CC sequence represents a fragment of a H. pylori 26 kDa protein-binding
 CC antibody heavy chain complementarity determining region CDR3 which is
 CC used to illustrate the method of the invention.

SO Sequence 13 AA;

Query Match 100.0%; Score 26; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FPFY 4

DB 5 PPFY 8

RESULT 6

ID AAB10071 standard; peptide: 13 AA.

XX AAB10071;

DT 03-NOV-2000 (first entry)

DE U1A protein derived tryptic peptide #10.

KW U1A protein; tryptic peptide; functional site; detection.

OS Unidentified.

PN WO200031544-A1.

PD 02-JUN-2000.

PF 23-NOV-1999; 99WO-EP09052.

PR 24-NOV-1998; 98DE-1054196.

PA (XERI-) XERION PHARM GMBH.

PI I1ag LL, Ng JH;

DR WPI; 2000-400193/34.

PT Identifying functional sites in proteins, useful for detecting epitopes and ligand binding sites, by complexing with tagged binding partner, then laser irradiation to alter the protein close to the tag.

PS Example 2; Page 24; 41pp; German.

CC This invention describes a novel method for identifying one or more functional sites in a target protein (I) which comprises: (i) complexing (I) with a binding partner (A) that carries a laser-activatable tag; (ii) irradiating the complex (II) formed with laser light to generate free radicals that alter bound (I) selectively at the binding sites; and (iii) identifying the altered region by a combination of protein cleavage and mass spectrometry (MS). The method is used to identify any type of functional site in any type of protein but particularly ligand-binding sites and epitopes. It can also be used to investigate pathologically altered proteins or oncogenic proteins, and to inactivate target pathophysiological proteins. The method does not require knowledge of the three-dimensional structure of (I), and is simple, rapid and automatable. CC It can identify non-linear or discontinuous epitopes, and determines protein function without inactivation of (I) (since the tag is bound to (A) not to (I) itself). The method may be used in vitro or in vivo for studying intra- or extra-cellular proteins. This sequence represents a U1A protein derived tryptic peptide which is used to illustrate the method of the invention.

SQ Sequence 13 AA;

Query Match 100.0%; Score 26; DB 21; Length 13;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4

DB 5 PPFY 8

RESULT 7

AAE02857

ID AAE02857 standard; peptide: 13 AA.

XX AAE02857;

XX 10-AUG-2001 (first entry)

DE Tryptic peptide #1 of human U2 small nuclear ribonucleoprotein B.

KW Human: U2 snRNP B; cytosolic; anti-sense therapy; tryptic peptide;

KW binding protein-based therapy; breast cancer-associated protein marker;

KW metastasis; U2 small nuclear ribonucleoprotein B; proteolysis; trypsin;

OS surface-enhanced laser desorption and ionisation; SELDI.

PN Homo sapiens.

PD WO200136470-A2.

PF 16-NOV-2000; 2000WO-US31483.

PR 16-NOV-1999; 99US-0165673.

PR 17-DEC-1999; 99US-0172170.

PR 27-JAN-2000; 2000US-0178860.

PR 03-MAY-2000; 2000US-0201721.

PR 10-NOV-2000; 2000US-0172170.

PA (MATR-) MATRITTECH INC.

PI Watkins B;

DR WPI; 2001-355615/37.

PT Novel breast cancer-associated proteins and nucleic acids encoding the proteins useful for diagnosing, treating breast cancer and as

PS Claim 43; Page 41; 64pp; English.

CC The present invention relates to an isolated, breast cancer-associated polypeptide detectable at a higher concentration in serum of a human having a characteristic molecular weight and sequence. The breast cancer-associated protein is used in an anti-sense or binding protein-based therapy. The breast cancer-associated protein is useful for diagnosing cancer, in particular breast cancer in an individual, by detecting the protein in breast tissue, blood, serum, plasma, sweat, tears, urine, peritoneal fluid, lymph, vaginal secretions, semen, spinal fluid, ascitic fluid, saliva, sputum or breast exudate isolated from the individual. The breast cancer associated proteins permit a rapid detection, preferably before metastasis occurs, of breast cancer. CC The present sequence is a tryptic peptide of human U2 small nuclear ribonucleoprotein B (U2 snRNP B). U2 snRNP B is a breast cancer-associated marker protein which has binding affinity to a nickel surface-enhanced laser desorption and ionisation (SELDI) chip. The peptide is obtained from U2 snRNP B by proteolysis with trypsin.

SQ Sequence 13 AA;

Query Match 100.0%; Score 26; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4

DB 5 PPFY 8

RESULT 8

AAE02910

ID AAE02910 standard; peptide: 15 AA.

XX AAE02910;

DT 10-AUG-2001 (first entry)

XX	DE	Tryptic peptide #1 of human U2 small nuclear ribonucleoprotein B.
KW	Human; marker molecule; therapy; cancer; breast; lung; prostate; bladder;	
KW	cervical; ovarian; colon; colorectal; tryptic peptide; proteolysis;	
XX	U2 small nuclear ribonucleoprotein B; U2 snRNP B; trypsin.	
OS	Homo sapiens.	
XX		
PN	WO200136977-A2.	
PD		
XX	25-MAY-2001.	
XX		
XX	16-NOV-2000; 2000WO-US31492.	
PR	16-NOV-1999; 99US-0165673.	
PR	17-DEC-1999; 99US-0172170.	
PR	27-JAN-2000; 2000US-0178860.	
PR	03-MAY-2000; 2000US-0201721.	
PR	10-NOV-2000; 2000US-0172170.	
XX		
PA	(MATR-) MATRITECH INC.	
XX		
PI	Watkins BA;	
XX		
DR	WPI; 2001-355675/37.	
XX		
PT	Identification of disease markers in mammals, especially cancer markers	
PT	useful as targets in methods for detecting or treating cancer, by	
PT	removing an abundant protein from a sample and analyzing by mass	
PT	spectroscopy based methods	
XX		
PS	Example 2; Page 26; 46pp; English.	
XX		
CC	The present invention relates to marker molecules indicative of disease	
CC	in mammals which are identified using a new method in which an abundant	
CC	protein is removed from a sample to enable less abundant molecules to be	
CC	evaluated as possible markers; the sample is then fractionated and	
CC	fractions analysed by known mass spectroscopy techniques and results	
CC	compared between individuals with and without the disease. The method is	
CC	useful for identifying disease markers in mammals, especially humans,	
CC	especially markers for cancer, such as breast, lung, prostate, bladder,	
CC	cervical, ovarian, colon or colorectal cancer. Once identified, the	
CC	markers can be used as targets in assays for detecting the disease and	
CC	in disease treatment e.g. administration of antibodies binding to and	
CC	eliminating/reducing activity of the target protein in vivo.	
CC	The present sequence is a tryptic peptide of human U2 small nuclear	
CC	ribonucleoprotein B (U2 snRNP B) which is used in the invention. The	
CC	peptide is obtained from U2 snRNP B by proteolysis with trypsin.	
CC	Note: The present sequence is also shown in sequence listing of the	
CC	specification, but lacks an amino acid in the N-terminal end and in the	
CC	C-terminal end.	
XX		
XX	Sequence 15 AA:	
XX		
QY	1 FPFY 4	
DB	6 FPFY 9	
XX		
XX	Query Match 100.0%; Score 26; DB 22; Length 15;	
XX	Best Local Similarity 100.0%; Prcd. NO. 62;	
XX	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX		
XX	RESULT 9	
XX	AA026562	
ID	AA026562 standard; Peptide: 15 AA.	
XX		
AC	AA026562;	
XX		
XX	06-MAR-2003 (first entry)	
DT		
DE	Theonline synthebase 13.75 protein N-terminal peptide region.	
XX		

```

KW Threonine synthetase 13.75 protein; enzyme; DNA recombination; cancer;
KM HIV infection.
OS Unidentified.
PN CN1342760-A.
XX
PD 03-APR-2002.
XX
PF 12-SEP-2000; 2000CN-0125129.
XX
PR 12-SEP-2000; 2000CN-0125129.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
PI Mao Y, xie Y;
XX
DR WPI; 2002-529801/57.
XX
PT Polypeptide-threonine synthetase 13.75 and polynucleotide for coding it
PT
XX
PS Example 5; Page 18 (Disclosure); 32pp; Chinese.
XX
CC The invention relates to a novel threonine synthetase 13.75 protein. The
CC invention also relates to the polynucleotide for coding it, the process
CC for preparing the polypeptide by DNA recombination technique, the
CC application of the polypeptide in treating several diseases such as
CC cancer, HIV infection, the antagonist against this polypeptide and its
CC therapeutic action, and the application of said polynucleotide to coding
CC this new threonine synthetase 13.75. This sequence represents an N-
CC terminal peptide region of the threonine synthetase 13.75 protein of the
CC invention.
XX
SQ Sequence 15 AA;
OY
Query Match 100.0%; Score 26; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
DB 10 FPFY 13
OY 1 FPFY 4
| | | |
DB 10 FPFY 13
RESULT 10
ABP46251
ID ABP46251 standard; peptide; 16 AA.
XX
AC ABP46251;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human Blys binding scfv VH CDR3 SEQ ID 2262.
XX
BLys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KM tumour necrosis factor; B cell proliferation; B cell differentiation;
KM immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KM antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KM systemic lupus erythematosus; rheumatoid arthritis; CYID; AIDS;
KM common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PF 15-JUN-2001; 2001WO-US19110.
XX
PR 16-JUN-2000; 2000US-212210P.
PR 17-OCT-2000; 2000US-240816P.
PR 16-MAR-2001; 2001US-276248P.
PR

```

PR 21-MAR-2001; 2001US-277379P.
 PR 25-MAY-2001; 2001US-293499P.
 XX
 PA (HUMA-) HUMAN GENOME SCI. INC.
 XX (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 DR
 XX
 PT Antibodies against B lymphocyte Stimulating polypeptides, useful for
 PT the diagnosis and treatment of cancers and immune disorders -
 PS Claim 2; Page 2960; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytototoxic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antineumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method
 CC of the invention.
 CC
 SQ Sequence 16 AA:
 XX
 QY
 DB 1 PFY 4
 11 PFY 14
 Query Match 100.0%; Score 26; DB 23; Length 16;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
 AAM48100
 ID AAM48100 standard; peptide; 20 AA.
 XX
 AC AAM48100;
 XX
 DT 15-JUN-1998 (first entry)
 DE NE-dlg SEQ ID NO:6 cancer related molecule from J1066581.
 XX
 KW Human; NE-dlg: discs large 1 gene; cancer related molecule; nerve;
 XX internal secretion tissue.
 OS Homo sapiens.
 XX
 PN JPI0066581-A.
 PD 10-MAR-1998.
 XX
 PF 23-AUG-1996; 96JP-0241370.
 XX
 PR 23-AUG-1996; 96JP-0241370.
 PA (SUME) SUMITOMO ELECTRIC IND CO.
 DR WPI; 1998-224339/20.
 XX
 PT Human discs large 1 gene family - useful in, e.g. therapeutic
 PT composition(s) for treating cancer

XX
 PS Claim 6; Page 22; 31pp; Japanese.
 CC
 CC The present sequence represents a NE-dlg molecule. The present invention
 CC describes human discs large 1 gene (dlg) family expressible in nerve
 CC tissue. Also described are: (1) a polynucleotide (PN) encoding dlg and
 CC comprising a 3100 (I) or 3213 bp (II) sequence; (2) RNA corresponding to
 CC PN having sequences (I) and (II); (3) RNA having a base sequence
 CC translated by dlg; (4) an antisense PN having a at least 15 bp sequence,
 CC and which is a part of PN of (1); (5) a derivative of the antisense PN;
 CC (6) an antibody specific for dlg, and (7) an antibody specific for a
 CC polypeptide having a 817 or 849 aa sequence. The polypeptide, RNA's and
 CC antibodies can be used for detection of dlg. The antisense PN can be
 CC used as a therapeutic composition for treating cancer.
 XX
 SQ Sequence 20 AA:
 XX
 QY
 DB 1 PFY 4
 11 PFY 19
 Query Match 100.0%; Score 26; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
 AAG77700
 ID AAG77700 standard; Protein; 22 AA.
 XX
 AC AAG77700;
 XX
 DT 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:8466.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 XX colorectal carcinoma.
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 XX
 PR 03-NOV-1999; 99US-0163280.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 DR WPI; 2001-235357/24.
 DR N-PSDB; AAH37107.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PS Claim 11; Page 9724; 9803pp; English.
 XX
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,

CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 CC
 XX
 SQ Sequence 22 AA:
 Query Match 100.0%; Score 26; DB 22; Length 22;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PPFY 4
 1111
 Db 12 PPFY 15
 RESULT 13
 AAB66072
 ID AAB66072 standard; Peptide: 23 AA.
 XX
 AC AAB66072;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human INTERCEPT 296 extracellular domain #1.
 XX
 KW TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.
 KW
 XX
 OS Homo sapiens.
 XX
 PN MO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14858.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 DR WPI: 2001-032313/04.
 XX
 PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Page 331; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 CC

SQ Sequence 23 AA:
 Query Match 100.0%; Score 26; DB 22; Length 23;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PPFY 4
 1111
 Db 8 PPFY 11
 RESULT 14
 AAW19473
 ID AAW19473 standard; peptide: 27 AA.
 XX
 AC AAW19473;
 XX
 DT 05-SEP-1997 (first entry)
 XX
 DE Anti-fungal and anti-bacterial D-amino acid histatin 2.
 XX
 KW Candida albicans; periodontitis; caries; tooth decay; oral infection;
 KW vaginal infection; urethral infection; mucosal infection; ear infection;
 KW respiratory infection; skin infection; ophthalmic infection;
 KW bacterial disease; Streptococcus mutans; Porphyromonas gingivalis;
 KW Clostridium histolyticum; human.
 KW
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..27
 FT /note= "At least one amino acid must have
 FT D-configuration"
 XX
 PN MO9640770-A2.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US09962.
 XX
 PR 07-JUN-1995; 95US-0485273.
 XX
 PA (PERT-) PERIODONTIX INC.
 XX
 PI (UYBO-) UNIV BOSTON.
 XX
 PI Friden PM, Oppenheim FG, Roberts FD, Spaccapoli P;
 XX
 DR WPI: 1997-052234/05.
 XX
 PS Anti-fungal and anti-bacterial D-amino acid histatins and
 PS histatin-based peptide(s) - useful eep for treating Candidal
 PS infections, periodontitis and caries.
 XX
 PT Disclosure: Fig 1; 63pp; English.
 XX
 CC The present sequence represents an anti-fungal and anti-bacterial
 CC D-amino acid histatin 2, found in human paratoid secretion, where at
 CC least one amino acid in the peptide is in the D-configuration. The
 CC peptide is preferably modified by an acetyl or carbamyl addition at
 CC the N-terminus and/or amidation at the C-terminus. The novel D-amino
 CC acid-containing peptide, based on the naturally occurring histidine-
 CC rich human histatins, have anti-fungal and anti-bacterial activity and
 CC are useful in compositions for the treatment of oral, vaginal, urethral,
 CC mucosal, respiratory, skin, ear and ophthalmic fungal or bacterial
 CC infections. It is particularly effective against local and systemic
 CC Candida albicans infection, against oral bacterial diseases such as
 CC caries and periodontitis (by inhibiting e.g. Streptococcus mutans and
 CC Porphyromonas gingivalis) and against Clostridium histolyticum. The D-
 CC amino acid-containing peptide has superior anti-fungal (especially
 CC anti-candidal) and anti-bacterial activity, particularly on a weight
 CC basis, compared to the natural L-amino acid forms of histatins and
 CC histatin-based peptides. The presence of D-residues also makes the

CC peptides more resistant to degradation than corresponding L-amino acid
 CC versions.
 XX
 SQ Sequence 27 AA;

Query Match 100.0%; Score 26; DB 18; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 Db 13 FPFY 16

RESULT 15
 AAM19428
 ID AAM19428 standard; peptide; 27 AA.

AC AAM19428;

DT 05-SEP-1997 (first entry)

DE Anti-fungal and anti-bacterial histatin-based histatin 2.

XX
 KW Candida albicans; periodontitis; caries; tooth decay; oral infection;
 KW vaginal infection; urethral infection; mucosal infection; ear infection;
 KW respiratory infection; skin infection; ophthalmic infection;
 KW bacterial disease; Streptococcus mutans; Porphyromonas gingivalis;
 KW Clostridium histolyticum; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 1..27
 FT Region /note= "At least one amino acid must have a
 D-configuration"

PN WO9640768-A2.

PD 19-DEC-1996.

PF 07-JUN-1996; 96MO-US09374.

PR 07-JUN-1995; 95US-0481888.

PR 01-NOV-1991; 91US-0786571.

PR 28-OCT-1993; 93US-0145030.

PR 09-AUG-1994; 94US-0287717.

PA (PERI-) PERIODONTIX INC.

PI (UYBO-) UNIV BOSTON.

PI Friden PM, Oppenheim FG, Roberts FD, Spacchiapoli P;

PT Anti-fungal and anti-bacterial histatin-based peptide(s) - useful
 PT esp for treating Candidal infections, periodontitis and caries.

PS Disclosure: Fig 1; 72pp; English.

XX The present sequence represents histatin 2, found in human paratoid
 CC secretion, where at least one amino acid is in the D-configuration.
 CC The peptide preferably is modified by an acetyl or carboxyl addition
 CC at the N-terminus and/or amidation at the C-terminus. The peptide,
 CC based on the naturally occurring histidine-rich human histatins, have
 CC anti-fungal and anti-bacterial activity and are useful in compositions
 CC for the treatment of oral, vaginal, urethral, mucosal, respiratory,
 CC skin, ear and ophthalmic fungal or bacterial infections. It is
 CC particularly effective against local and systemic Candida albicans
 CC infection, against oral bacterial diseases such as caries and
 CC periodontitis (by inhibiting e.g. Streptococcus mutans and Porphyromonas
 CC gingivalis) and against Clostridium histolyticum. The peptide has

CC superior anti-fungal (especially anti-Candidal) and anti-bacterial
 CC activity, particularly on a weight basis, compared to the longer,
 CC naturally occurring histatins. Peptides containing D-residues are also
 CC more resistant to degradation than L-amino acid versions.

SQ Sequence 27 AA;

Query Match 100.0%; Score 26; DB 18; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
 ||||
 Db 13 FPFY 16

Search completed: August 20, 2003, 12:33:45
 Job time : 10.7952 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 2.77108 Seconds
(without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-31
Perfect score: 26
Sequence: 1 FPPY 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCYUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	6	1	US-08-287-717-14
2	26	100.0	6	1	US-08-441-914-14
3	26	100.0	27	1	US-08-481-888A-2
4	26	100.0	27	1	US-08-485-273A-2
5	26	100.0	27	2	US-08-973-559A-2
6	26	100.0	27	2	US-08-973-559-2
7	26	100.0	27	4	US-09-693-822B-29
8	26	100.0	38	1	US-08-287-717-9
9	26	100.0	38	1	US-08-481-888A-1
10	26	100.0	38	1	US-08-485-273A-1
11	26	100.0	38	1	US-08-441-914-9
12	26	100.0	38	2	US-08-973-559A-1
13	26	100.0	38	2	US-08-973-559-1
14	26	100.0	38	4	US-09-693-822B-28
15	26	100.0	39	1	US-08-189-331-143
16	26	100.0	39	2	US-08-471-068-143
17	26	100.0	47	4	US-09-904-615-88
18	26	100.0	53	4	US-09-149-476-439
19	26	100.0	61	4	US-09-107-532A-7043
20	26	100.0	66	4	US-09-134-001C-4702
21	26	100.0	76	3	US-09-083-351-19
22	26	100.0	76	3	US-09-083-352-19
23	26	100.0	82	4	US-09-107-532A-6005
24	26	100.0	98	1	US-07-906-930E-5
25	26	100.0	106	3	US-09-083-351-4
26	26	100.0	106	3	US-09-083-351-5
27	26	100.0	106	3	US-09-083-351-6

28	26	100.0	106	3	US-09-083-351-8	Sequence 8, Appl
29	26	100.0	106	3	US-09-083-351-10	Sequence 10, Appl
30	26	100.0	106	3	US-09-083-352-4	Sequence 4, Appl
31	26	100.0	106	3	US-09-083-352-5	Sequence 5, Appl
32	26	100.0	106	3	US-09-083-352-6	Sequence 6, Appl
33	26	100.0	106	3	US-09-083-352-8	Sequence 8, Appl
34	26	100.0	106	3	US-09-083-352-10	Sequence 10, Appl
35	26	100.0	118	4	US-09-732-210-675	Sequence 675, App
36	26	100.0	177	4	US-09-149-476-565	Sequence 565, App
37	26	100.0	184	2	US-08-865-336-1	Sequence 1, Appl
38	26	100.0	184	1	US-07-741-940-6	Sequence 6, Appl
39	26	100.0	185	1	US-08-289-548A-6	Sequence 6, Appl
40	26	100.0	185	1	US-08-452-654-6	Sequence 6, Appl
41	26	100.0	185	1	US-08-452-655B-6	Sequence 6, Appl
42	26	100.0	185	2	US-08-865-336-4	Sequence 4, Appl
43	26	100.0	185	3	US-08-450-582-6	Sequence 6, Appl
44	26	100.0	185	4	US-08-449-731-6	Sequence 6, Appl
45	26	100.0	192	4	US-09-199-637A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-287-717-14
; Sequence 14, Application US/08287717
; Patent No. 5486503
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; TITLE OF INVENTION: No. 5486503el Antl-Fungal Peptides and Uses
; NUMBER OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,717
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,030
; FILING DATE:
; APPLICATION NUMBER: US 07/786,571
; FILING DATE: 01-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: B091-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-287-717-14
; Query Match 100.0%; Score 26; DB 1; Length 6;
; Best Local Similarity 100.0%; Pred. No. 2.5e+05;
; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPY 4
|||
Db 1 PPY 4

RESULT 2

US-08-441-914-14
; Sequence 14, Application US/08441914
; Patent No. 5696078
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; TITLE OF INVENTION: No. 5696078e1 Anti-Fungal Peptides and Uses
; TITLE OF INVENTION: Theof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,914
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,717
; FILING DATE: 09-AUG-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,030
; FILING DATE: 28-OCT-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786,571
; FILING DATE: 01-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: David E. Brook
; REGISTRATION NUMBER: 22,592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-441-914-14

Query Match 100.0%; Score 26; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPY 4
|||
Db 1 PPY 4

RESULT 3

US-08-481-888A-2
; Sequence 2, Application US/08481888A
; Patent No. 5631228
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; APPLICANT: Roberts, F. Donald
; TITLE OF INVENTION: ANTI-FUNGAL AND ANTI-BACTERIAL

;; TITLE OF INVENTION: HISTATIN-BASED PEPTIDES
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: MA
;; COUNTRY: US
;; ZIP: 02173
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/481,888A
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/287,717
;; FILING DATE: 09-AUG-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/145,030
;; FILING DATE: 28-OCT-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/786,571
;; FILING DATE: 01-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brook, David E.
;; REGISTRATION NUMBER: 22,592
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 617-861-9540
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-481-888A-2

Query Match 100.0%; Score 26; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPY 4
|||
Db 13 PPY 16

RESULT 4

US-08-485-273A-2
; Sequence 2, Application US/08485273A
; Patent No. 5646119
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; APPLICANT: Spaccapoli, Peter
; TITLE OF INVENTION: D-Amino Acid Histatin-Based Peptides as
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; SOFTWARE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,273A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,717
FILING DATE: 09-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/145,030
FILING DATE: 28-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/786,571
FILING DATE: 01-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-02A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 1..27
OTHER INFORMATION: /note="At least one amino acid
US-08-485-273A-2
OTHER INFORMATION: must have a D configuration"

Query Match 100.0%; Score 26; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
||||
DB 13 PPFY 16

RESULT 5
US-08-973-563A-2
Sequence 2, Application US/08973563A
Patent No. 5885965
GENERAL INFORMATION:
APPLICANT: Oppenheim, Frank G.
APPLICANT: Xu, Tao
APPLICANT: Spacciapoli, Peter
APPLICANT: Roberts, F. D.
APPLICANT: Friden, Philip M.
TITLE OF INVENTION: Anti-Fungal D-Amino Acid Histatin-Based
TITLE OF INVENTION: Peptides
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,563A
FILING DATE: 07-JUN-1996

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,273
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-02A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 1..27
OTHER INFORMATION: /note="At least one amino acid
US-08-973-563A-2
OTHER INFORMATION: must have a D configuration"

Query Match 100.0%; Score 26; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
||||
DB 13 PPFY 16

RESULT 6
US-08-973-559-2
Sequence 2, Application US/08973559
Patent No. 5912230
GENERAL INFORMATION:
APPLICANT: OPPENHEIM, FRANK G.
APPLICANT: XU, TAO
APPLICANT: ROBERTS, F. D.
APPLICANT: SPACCIAPOLI, PETER
APPLICANT: FRIDEN, PHILIP M.
TITLE OF INVENTION: Anti-Fungal and Anti-Bacterial
TITLE OF INVENTION: Histatin-Based Peptides
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,559
FILING DATE: 07-JUN-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,888
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-01A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540

;; INFORMATION FOR SEQ ID NO: 2;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-973-559-2

Query Match 100.0%; Score 26; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
1111
DB 13 FFY 16

RESULT 7
US-09-693-822B-29
; Sequence 29, Application US/09693822B
; Patent No. 6555650
; GENERAL INFORMATION:
; APPLICANT: Lajoie, Gilles A.
; TITLE OF INVENTION: Cyclic Analogs of Histatins
; FILE REFERENCE: 36555-0002
; CURRENT APPLICATION NUMBER: US/09/693,822B
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: CA 2,285,673
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-693-822B-29

Query Match 100.0%; Score 26; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
1111
DB 13 FFY 16

RESULT 8
US-08-287-717-9
; Sequence 9, Application US/08287717
; Patent No. 5486503
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; TITLE OF INVENTION: No. 5486503el Anti-Fungal Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,717
; FILING DATE:
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/145,030
;; FILING DATE:
;; APPLICATION NUMBER: US 07/786,571
;; FILING DATE: 01-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: B091-17
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 617-861-9540
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-287-717-9

Query Match 100.0%; Score 26; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
1111
DB 24 FFY 27

RESULT 9
US-08-481-888A-1
; Sequence 1, Application US/08481888A
; Patent No. 5631228
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; APPLICANT: Roberts, F. Donald
; TITLE OF INVENTION: ANTI-FUNGAL AND ANTI-BACTERIAL
; TITLE OF INVENTION: HISTATIN-BASED PEPTIDES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,888A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/287,717
; FILING DATE: 09-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,030
; FILING DATE: 28-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/786,571
; FILING DATE: 01-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: PER95-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /product= "PSE"
US-08-481-888A-1

Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 38;
Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 PPY 4
DB 24 PPY 27
```

```
RESULT 10
US-08-485-273A-1
Sequence 1, Application US/08485273A
Patent No. 5646119
GENERAL INFORMATION:
APPLICANT: Oppenheim, Frank G.
APPLICANT: Xu, Tao
TITLE OF INVENTION: D-Amino Acid Histatin-Based Peptides as
Title of Invention: Anti-Fungal and Anti-Bacterial Agents
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,273A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,717
FILING DATE: 09-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/145,030
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US 07/786,571
FILING DATE: 01-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-02A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
```

```
LOCATION: 2
OTHER INFORMATION: /product= "PSE"
FEATURE:
NAME/KEY: Region
LOCATION: 1..38
OTHER INFORMATION: /note= "At least one amino acid
OTHER INFORMATION: must have a D configuration."
US-08-485-273A-1

Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 38;
Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 PPY 4
DB 24 PPY 27
```

```
RESULT 11
US-08-441-914-9
Sequence 9, Application US/08441914
Patent No. 5696078
GENERAL INFORMATION:
APPLICANT: Oppenheim, Frank G.
APPLICANT: Xu, Tao
TITLE OF INVENTION: No. 5696078el Anti-Fungal Peptides and Uses
Title of Invention: Thereof
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,914
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/287,717
FILING DATE: 09-AUG-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/145,030
FILING DATE: 28-OCT-93
APPLICATION NUMBER: 07/786,571
FILING DATE: 01-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: David E. Brook
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: B091-17F22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-914-9
```

```
Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 38;
Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 PPY 4
```

DB 24 PFY 27

```
|||||
US-08-973-563A-1
; Sequence 1, Application US/08973563A
; Patent No. 5885965
; GENERAL INFORMATION:
; APPLICANT: Oppenheim, Frank G.
; APPLICANT: Xu, Tao
; APPLICANT: Spacciapoli, Peter
; APPLICANT: Roberts, F. D.
; APPLICANT: Friden, Philip M.
; TITLE OF INVENTION: Anti-Fungal D-Amino Acid Histatin-Based
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973.563A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,273
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: PER95-02A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /product="PSE"
; NAME/KEY: Region
; LOCATION: 1..38
; OTHER INFORMATION: /note="At least one amino acid
; OTHER INFORMATION: must have a D configuration."
US-08-973-563A-1

Query Match 100.0%; Score 26; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PFY 4
|||||
DB 24 PFY 27

RESULT 13
US-08-973-559-1
; Sequence 1, Application US/08973559
```

```
;; Patent No. 5912230
;; GENERAL INFORMATION:
;; APPLICANT: OPPENHEIM, FRANK G.
;; APPLICANT: XU, TAO
;; APPLICANT: ROBERTS, F. D.
;; APPLICANT: SPACCIAPOLI, PETER
;; APPLICANT: FRIDEN, PHILIP M.
;; TITLE OF INVENTION: Anti-Fungal and Anti-Bacterial
;; TITLE OF INVENTION: Histatin-Based Peptides
;; NUMBER OF SEQUENCES: 37
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: MA
;; COUNTRY: US
;; ZIP: 02173
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/973,559
;; FILING DATE: 07-JUN-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/481,888
;; FILING DATE: 07-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brook, David E.
;; REGISTRATION NUMBER: 22,592
;; REFERENCE/DOCKET NUMBER: PER95-01A2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 781-861-6240
;; TELEFAX: 781-861-9540
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 2
;; OTHER INFORMATION: /note="*/product="PSE""
US-08-973-559-1

Query Match 100.0%; Score 26; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PFY 4
|||||
DB 24 PFY 27

RESULT 14
US-09-693-822B-28
; Sequence 28, Application US/09693822B
; Patent No. 6555650
; GENERAL INFORMATION:
; APPLICANT: Lajotte, Gilles A.
; FILE REFERENCE: 36555-0002
; CURRENT APPLICATION NUMBER: US/09/693,822B
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: CA 2,285,673
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 28
```

LENGTH: 38
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (2)..(2)
OTHER INFORMATION: phosphoserine
US-09-693-822B-28

Query Match 100.0%; Score 26; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
DB 24 FPFY 27

RESULT 15
US-08-189-331-143
Sequence 143 Application US/08189331
Patent No. 5747334
GENERAL INFORMATION:

APPLICANT: Kay, B. K.
APPLICANT: Fowles, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/189,331
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 143:

SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-189-331-143

Query Match 100.0%; Score 26; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
DB 27 FPFY 30

Search completed: August 20, 2003, 12:44:21
Job time: 3.77108 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 4.33735 Seconds

(Without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26
Sequence: 1 PFY 4Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 100%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PC7_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	4	11	US-09-300-425B-31
2	26	100.0	10	10	US-09-912-672A-10
3	26	100.0	16	11	US-09-880-748-2262
4	26	100.0	22	15	US-10-106-698-8476
5	26	100.0	23	11	US-09-759-1308-426
6	26	100.0	23	14	US-10-042-431-56
7	26	100.0	39	9	US-09-864-761-46278
8	26	100.0	47	9	US-09-739-254-88
9	26	100.0	47	9	US-09-904-615-88
10	26	100.0	47	12	US-10-055-098-88
11	26	100.0	47	15	US-10-054-988-88
12	26	100.0	48	11	US-09-895-298-161
13	26	100.0	53	11	US-09-809-391-439
14	26	100.0	68	9	US-09-764-887-237
15	26	100.0	68	15	US-10-073-961-237

16	26	100.0	70	11	US-09-764-891-4300	Sequence 4300, App
17	26	100.0	78	10	US-09-764-877-1056	Sequence 1056, App
18	26	100.0	89	9	US-09-925-301-1058	Sequence 1058, App
19	26	100.0	89	15	US-10-106-698-5417	Sequence 5417, App
20	26	100.0	93	10	US-09-895-494-108	Sequence 108, App
21	26	100.0	98	15	US-10-011-585A-176	Sequence 176, App
22	26	100.0	117	15	US-10-036-542-61	Sequence 61, App
23	26	100.0	118	9	US-09-815-242-11401	Sequence 11401, App
24	26	100.0	118	9	US-09-815-242-11560	Sequence 11560, App
25	26	100.0	118	9	US-09-867-550-198	Sequence 198, App
26	26	100.0	146	9	US-09-811-284-188	Sequence 188, App
27	26	100.0	147	10	US-09-764-847-972	Sequence 972, App
28	26	100.0	147	15	US-10-092-154-972	Sequence 972, App
29	26	100.0	177	11	US-09-809-391-565	Sequence 565, App
30	26	100.0	177	11	US-09-888-0678-62	Sequence 62, App
31	26	100.0	177	15	US-10-013-315-38	Sequence 38, App
32	26	100.0	189	9	US-09-765-205-4	Sequence 4, App
33	26	100.0	189	10	US-09-731-872-356	Sequence 356, App
34	26	100.0	189	11	US-09-946-374-8	Sequence 8, App
35	26	100.0	189	12	US-10-015-387A-8	Sequence 8, App
36	26	100.0	189	12	US-10-006-130A-8	Sequence 8, App
37	26	100.0	189	12	US-10-199-672-126	Sequence 126, App
38	26	100.0	189	12	US-09-876-997-356	Sequence 356, App
39	26	100.0	189	12	US-10-006-172A-8	Sequence 8, App
40	26	100.0	189	12	US-10-187-749-126	Sequence 126, App
41	26	100.0	189	12	US-10-194-457-126	Sequence 126, App
42	26	100.0	189	14	US-10-052-586-126	Sequence 126, App
43	26	100.0	189	15	US-10-174-590-126	Sequence 126, App
44	26	100.0	189	15	US-10-176-758-126	Sequence 126, App
45	26	100.0	189	15	US-10-175-737-126	Sequence 126, App

ALIGNMENTS

RESULT 1
US-09-300-425B-31
Sequence 31, Application US/09300425B
Publication No. US20030045681A1
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARKI, Lorenzo
APPLICANT: VITTI, Francesca
APPLICANT: BICHER, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
TITLE OF INVENTION: ANGIOGENESIS
FILE REFERENCE: SCH-1733P1
CURRENT APPLICATION NUMBER: US/09/300,425B
CURRENT FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 09/075,338
PRIOR FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: anti-BD-B
OTHER INFORMATION: antibody clone
US-09-300-425B-31
Query Match
Best Local Similarity 100.0%; Score 26; DB 11; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PFY 4
Db 1 PFY 4

RESULT 2

US-09-912-672A-10
; Sequence 10, Application US/09912672A
; Patent No. US2002016469A1
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: CLASS II CYTOKINE RECEPTOR-LIKE PROTEINS
; FILE REFERENCE: 07334-184001
; CURRENT APPLICATION NUMBER: US/09/912,672A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/475,541
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-912-672A-10

Query Match 100.0%; Score 26; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
11111
DB 3 FPFY 6

RESULT 3
US-09-880-748-2262
; Sequence 2262, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2262
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2262

Query Match 100.0%; Score 26; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
11111
DB 11 FPFY 14

RESULT 4
US-10-106-698-8476
; Sequence 8476, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 8476
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-8476

Query Match 100.0%; Score 26; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
11111
DB 12 FPFY 15

RESULT 5
US-09-759-130B-426
; Sequence 426, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Sharp, John D.
; APPLICANT: Barnes, Thomas S.
; APPLICANT: Kirst, Susan J.
; APPLICANT: Mackay, Charles R.
; APPLICANT: Myers, Paul S.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MP100-5350NM1
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19

NUMBER OF SEQ ID NOS: 460
SOFTWARE: fastseq for Windows Version 4.0
SEQ ID NO 426
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-1308-426

Query Match 100.0%; Score 26; DB 11; Length 23;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
1111
DB 8 FFY 11

RESULT 6
US-10-042-431-56
Sequence 56, Application US/10042431
Publication No. US20020182675A1
GENERAL INFORMATION:
APPLICANT: MCCARTHY, Sean A
APPLICANT: BARNES, Thomas M
APPLICANT: FRASER, Christopher C
APPLICANT: SHARP, John D
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES.
FILE REFERENCE: 10147-602
CURRENT APPLICATION NUMBER: US/10/042,431
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24
NUMBER OF SEQ ID NOS: 79
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 56
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-431-56

Query Match 100.0%; Score 26; DB 14; Length 23;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
1111
DB 8 FFY 11

RESULT 7
US-09-864-761-46278
Sequence 46278, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46278
LENGTH: 39
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO A1121981.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
US-09-864-761-46278

Query Match 100.0%; Score 26; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
1111
DB 22 FFY 25

RESULT 8
US-09-739-254-88
Sequence 88, Application US/09739254
Patent No. US20010021700A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
CURRENT APPLICATION NUMBER: US/09/739,254
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: PCT/US99/19330
PRIOR FILING DATE: 1999-08-24
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 88
LENGTH: 47

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals stop translation
US-09-739-254-88
```

```

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FPFY 4
Db 27 FPFY 30
```

```

RESULT 9
US-09-904-615-88
; Sequence 88, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 88
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals stop translation
US-09-904-615-88
```

```

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FPFY 4
Db 27 FPFY 30
```

```

RESULT 10
US-10-055-098-88
; Sequence 88, Application US/10055098
; Publication No. US20030139954A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/10/055,098
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-02-23
```

```

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/19330
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 88
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals stop translation
US-10-055-098-88
```

```

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FPFY 4
Db 27 FPFY 30
```

```

RESULT 11
US-10-054-988-88
; Sequence 88, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/10/054,988
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 88
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals stop translation
US-10-054-988-88
```

```

Query Match
Best Local Similarity 100.0%; Score 26; DB 15; Length 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FPFY 4
Db 27 FPFY 30
```

RESULT 12

US-09-895-298-161
; Sequence 161, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29950
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/112,809
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 161
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-298-161

Query Match 100.0%; Score 26; DB 11; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
DB 13 FPFY 16

RESULT 13
US-09-809-391-439
; Sequence 439, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (24)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-809-391-439

Query Match 100.0%; Score 26; DB 11; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
DB 30 FPFY 33

RESULT 14
US-09-764-887-237
; Sequence 237, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA113
; CURRENT APPLICATION NUMBER: US/09/764, 887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 237
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-237

Query Match 100.0%; Score 26; DB 9; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
DB 65 FPFY 68

RESULT 15
US-10-073-961-237
; Sequence 237, Application US/10073961
; Publication No. US20030077602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21

[illegible]

;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08

Query Match 100.0%; Score 26; DB 15; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PPFY 4
1111
Db 65 PPFY 68

Search completed: August 20, 2003, 13:16:46
Job time : 4.33735 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:35:06 ; Search time 40.9398 Seconds
(without alignments)
85.038 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26

Sequence: 1 PPRY 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCRTUS.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US097A.COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097B.COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099A.COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US099B.COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*
28: /cgn2_6/ptodata/1/paa/US103.COMB.pep:*
29: /cgn2_6/ptodata/1/paa/US104.COMB.pep:*
30: /cgn2_6/ptodata/1/paa/US104.COMB.pep:*
31: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	26	100.0	4	14	US-09-075-338C-31
2	26	100.0	4	17	US-09-300-425B-31

3	26	100.0	4	19	US-09-512-082-31	Sequence 31, Appl
4	26 <td>100.0</td> <td>9</td> <td>23</td> <td>US-09-856-285-9</td> <td>Sequence 9, Appl1</td>	100.0	9	23	US-09-856-285-9	Sequence 9, Appl1
5	26 <td>100.0</td> <td>10</td> <td>16</td> <td>US-09-924-669-10</td> <td>Sequence 10, Appl</td>	100.0	10	16	US-09-924-669-10	Sequence 10, Appl
6	26 <td>100.0</td> <td>10</td> <td>18</td> <td>US-09-475-541-10</td> <td>Sequence 10, Appl</td>	100.0	10	18	US-09-475-541-10	Sequence 10, Appl
7	26 <td>100.0</td> <td>10</td> <td>23</td> <td>US-09-856-285-10</td> <td>Sequence 10, Appl</td>	100.0	10	23	US-09-856-285-10	Sequence 10, Appl
8	26 <td>100.0</td> <td>10</td> <td>24</td> <td>US-09-912-672-10</td> <td>Sequence 10, Appl</td>	100.0	10	24	US-09-912-672-10	Sequence 10, Appl
9	26 <td>100.0</td> <td>11</td> <td>23</td> <td>US-09-912-672A-10</td> <td>Sequence 10, Appl</td>	100.0	11	23	US-09-912-672A-10	Sequence 10, Appl
10	26 <td>100.0</td> <td>11</td> <td>23</td> <td>US-09-856-285-11</td> <td>Sequence 11, Appl</td>	100.0	11	23	US-09-856-285-11	Sequence 11, Appl
11	26 <td>100.0</td> <td>12</td> <td>3</td> <td>US-07-867-819A-86</td> <td>Sequence 86, Appl</td>	100.0	12	3	US-07-867-819A-86	Sequence 86, Appl
12	26 <td>100.0</td> <td>12</td> <td>3</td> <td>US-07-867-819B-86</td> <td>Sequence 86, Appl</td>	100.0	12	3	US-07-867-819B-86	Sequence 86, Appl
13	26 <td>100.0</td> <td>12</td> <td>3</td> <td>US-07-867-819B-86</td> <td>Sequence 86, Appl</td>	100.0	12	3	US-07-867-819B-86	Sequence 86, Appl
14	26 <td>100.0</td> <td>12</td> <td>3</td> <td>US-07-867-819B-86</td> <td>Sequence 86, Appl</td>	100.0	12	3	US-07-867-819B-86	Sequence 86, Appl
15	26 <td>100.0</td> <td>12</td> <td>8</td> <td>US-08-475-955-86</td> <td>Sequence 86, Appl</td>	100.0	12	8	US-08-475-955-86	Sequence 86, Appl
16	26 <td>100.0</td> <td>13</td> <td>21</td> <td>US-09-856-285-12</td> <td>Sequence 12, Appl</td>	100.0	13	21	US-09-856-285-12	Sequence 12, Appl
17	26 <td>100.0</td> <td>13</td> <td>21</td> <td>US-09-709-947-1</td> <td>Sequence 1, Appl1</td>	100.0	13	21	US-09-709-947-1	Sequence 1, Appl1
18	26 <td>100.0</td> <td>13</td> <td>21</td> <td>US-09-709-947-1</td> <td>Sequence 1, Appl1</td>	100.0	13	21	US-09-709-947-1	Sequence 1, Appl1
19	26 <td>100.0</td> <td>13</td> <td>23</td> <td>US-09-842-776A-15</td> <td>Sequence 15, Appl</td>	100.0	13	23	US-09-842-776A-15	Sequence 15, Appl
20	26 <td>100.0</td> <td>13</td> <td>23</td> <td>US-09-856-285-13</td> <td>Sequence 13, Appl</td>	100.0	13	23	US-09-856-285-13	Sequence 13, Appl
21	26 <td>100.0</td> <td>16</td> <td>1</td> <td>PCT-US01-19110-2262</td> <td>Sequence 2262, Ap</td>	100.0	16	1	PCT-US01-19110-2262	Sequence 2262, Ap
22	26 <td>100.0</td> <td>16</td> <td>1</td> <td>PCT-US02-36496-2262</td> <td>Sequence 2262, Ap</td>	100.0	16	1	PCT-US02-36496-2262	Sequence 2262, Ap
23	26 <td>100.0</td> <td>16</td> <td>23</td> <td>US-09-880-748-2262</td> <td>Sequence 2262, Ap</td>	100.0	16	23	US-09-880-748-2262	Sequence 2262, Ap
24	26 <td>100.0</td> <td>16</td> <td>28</td> <td>US-10-293-418-2262</td> <td>Sequence 2, Ap</td>	100.0	16	28	US-10-293-418-2262	Sequence 2, Ap
25	26 <td>100.0</td> <td>20</td> <td>16</td> <td>US-09-359-412-6</td> <td>Sequence 6, Appl1</td>	100.0	20	16	US-09-359-412-6	Sequence 6, Appl1
26	26 <td>100.0</td> <td>20</td> <td>16</td> <td>US-09-259-412A-6</td> <td>Sequence 6, Appl1</td>	100.0	20	16	US-09-259-412A-6	Sequence 6, Appl1
27	26 <td>100.0</td> <td>22</td> <td>1</td> <td>PCT-US00-26524B-8466</td> <td>Sequence 8466, Ap</td>	100.0	22	1	PCT-US00-26524B-8466	Sequence 8466, Ap
28	26 <td>100.0</td> <td>22</td> <td>27</td> <td>US-10-106-698-8476</td> <td>Sequence 8476, Ap</td>	100.0	22	27	US-10-106-698-8476	Sequence 8476, Ap
29	26 <td>100.0</td> <td>23</td> <td>1</td> <td>PCT-US00-14858-56</td> <td>Sequence 56, Appl</td>	100.0	23	1	PCT-US00-14858-56	Sequence 56, Appl
30	26 <td>100.0</td> <td>23</td> <td>1</td> <td>PCT-US00-14858B-56</td> <td>Sequence 56, Appl</td>	100.0	23	1	PCT-US00-14858B-56	Sequence 56, Appl
31	26 <td>100.0</td> <td>23</td> <td>16</td> <td>US-09-209-662B-672</td> <td>Sequence 672, Ap</td>	100.0	23	16	US-09-209-662B-672	Sequence 672, Ap
32	26 <td>100.0</td> <td>23</td> <td>17</td> <td>US-09-333-159-56</td> <td>Sequence 56, Appl</td>	100.0	23	17	US-09-333-159-56	Sequence 56, Appl
33	26 <td>100.0</td> <td>23</td> <td>19</td> <td>US-09-578-063-56</td> <td>Sequence 56, Appl</td>	100.0	23	19	US-09-578-063-56	Sequence 56, Appl
34	26 <td>100.0</td> <td>23</td> <td>22</td> <td>US-09-759-130B-426</td> <td>Sequence 426, Appl</td>	100.0	23	22	US-09-759-130B-426	Sequence 426, Appl
35	26 <td>100.0</td> <td>23</td> <td>26</td> <td>US-10-042-431-56</td> <td>Sequence 56, Appl</td>	100.0	23	26	US-10-042-431-56	Sequence 56, Appl
36	26 <td>100.0</td> <td>23</td> <td>28</td> <td>US-10-219-793-672</td> <td>Sequence 672, Appl</td>	100.0	23	28	US-10-219-793-672	Sequence 672, Appl
37	26 <td>100.0</td> <td>23</td> <td>24</td> <td>US-60-160-6203-4675</td> <td>Sequence 4675, Ap</td>	100.0	23	24	US-60-160-6203-4675	Sequence 4675, Ap
38	26 <td>100.0</td> <td>31</td> <td>34</td> <td>US-09-935-625-14634</td> <td>Sequence 14634, A</td>	100.0	31	34	US-09-935-625-14634	Sequence 14634, A
39	26 <td>100.0</td> <td>31</td> <td>30</td> <td>US-10-405-927-4007</td> <td>Sequence 4007, Ap</td>	100.0	31	30	US-10-405-927-4007	Sequence 4007, Ap
40	26 <td>100.0</td> <td>34</td> <td>1</td> <td>PCT-US01-08656-8043</td> <td>Sequence 8043, Ap</td>	100.0	34	1	PCT-US01-08656-8043	Sequence 8043, Ap
41	26 <td>100.0</td> <td>36</td> <td>19</td> <td>US-09-513-996A-77180</td> <td>Sequence 77180, A</td>	100.0	36	19	US-09-513-996A-77180	Sequence 77180, A
42	26 <td>100.0</td> <td>36</td> <td>26</td> <td>US-10-091-007-34</td> <td>Sequence 34, Appl</td>	100.0	36	26	US-10-091-007-34	Sequence 34, Appl
43	26 <td>100.0</td> <td>37</td> <td>30</td> <td>US-10-424-599-266152</td> <td>Sequence 266152, A</td>	100.0	37	30	US-10-424-599-266152	Sequence 266152, A
44	26 <td>100.0</td> <td>37</td> <td>31</td> <td>US-60-452-680-17495</td> <td>Sequence 17495, A</td>	100.0	37	31	US-60-452-680-17495	Sequence 17495, A
45	26 <td>100.0</td> <td>37</td> <td>31</td> <td>US-60-453-050-10630</td> <td>Sequence 10630, A</td>	100.0	37	31	US-60-453-050-10630	Sequence 10630, A

ALIGNMENTS

RESULT 1
US-09-075-338C-31 : Application US/09075338C
; Sequence 31, Appl1
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITTI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
; FILE REFERENCE: SCH-1733
; CURRENT APPLICATION NUMBER: US/09/075,338C
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
; OTHER INFORMATION: antibody clone
US-09-075-338C-31

Query Match 100.0%; Score 26; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 5.1e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
||||
Db 1 FPFY 4

RESULT 2
US-09-300-425B-31
; Sequence 31, Application US/09300425B
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITTI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-300-425B-31

Query Match Best Local Similarity 100.0%; Score 26; DB 17; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
||||
Db 1 FPFY 4

RESULT 3
US-09-512-082-31
; Sequence 31, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITTI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; FILE REFERENCE: SCH-1733P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-512-082-31

Query Match 100.0%; Score 26; DB 19; Length 4;

Best Local Similarity 100.0%; Pred. No. 5.1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
||||
Db 1 FPFY 4

RESULT 4
US-09-856-285-9
; Sequence 9, Application US/09856285
; GENERAL INFORMATION:
; APPLICANT: Ilag, Leodevico L.
; APPLICANT: Ng, Jocelyn, H.
; TITLE OF INVENTION: Method for Modifying and Identifying
; TITLE OF INVENTION: Functional Sites in Proteins
; FILE REFERENCE: 50125/024001
; CURRENT APPLICATION NUMBER: US/09/856,285
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/EP99/09052
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: DE 19854196.1
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-856-285-9

Query Match Best Local Similarity 100.0%; Score 26; DB 23; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
||||
Db 1 FPFY 4

RESULT 5
US-09-224-669-10
; Sequence 10, Application US/09224669
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: CLASS II CYTOKINE RECEPTOR-LIKE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: 09404/069001
; CURRENT APPLICATION NUMBER: US/09/224,669
; CURRENT FILING DATE: 1998-12-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-224-669-10

Query Match Best Local Similarity 100.0%; Score 26; DB 16; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
||||
Db 3 FPFY 6

RESULT 6
US-09-475-541-10
; Sequence 10, Application US/09475541
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: CLASS II CYTOKINE RECEPTOR-LIKE PROTEINS
US-09-475-541-10


```

; TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 07334-184001
; CURRENT APPLICATION NUMBER: US/09/475,541
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/224,669
; PRIOR FILING DATE: 1998-12-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-475-541-10

```

```

Query Match          100.0%; Score 26; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FPFY 4
        ||||
Db      3 FPFY 6

```

```

RESULT 7
US-09-856-285-10
; Sequence 10, Application US/09856285
; GENERAL INFORMATION:
; APPLICANT: Ila9, Leodevico L.
; APPLICANT: Ng, Jocelyn, H.
; TITLE OF INVENTION: Method for Modifying and Identifying
; FILE REFERENCE: 50125/024001
; CURRENT APPLICATION NUMBER: US/09/856,285
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/EP99/09052
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: DE 19854196.1
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-856-285-10

```

```

Query Match          100.0%; Score 26; DB 23; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FPFY 4
        ||||
Db      2 FPFY 5

```

```

RESULT 8
US-09-912-672-10
; Sequence 10, Application US/09912672
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: CLASS II CYTOKINE RECEPTOR-LIKE PROTEINS
; TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 07334-184001
; CURRENT APPLICATION NUMBER: US/09/912,672
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/475,541
; PRIOR FILING DATE: 1998-12-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

US-09-912-672-10

```

```

Query Match          100.0%; Score 26; DB 24; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FPFY 4
        ||||
Db      3 FPFY 6

```

```

RESULT 9
US-09-912-672A-10
; Sequence 10, Application US/09912672A
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: CLASS II CYTOKINE RECEPTOR-LIKE PROTEINS
; TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 07334-184001
; CURRENT APPLICATION NUMBER: US/09/912,672A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/475,541
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-912-672A-10

```

```

Query Match          100.0%; Score 26; DB 24; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FPFY 4
        ||||
Db      3 FPFY 6

```

```

RESULT 10
US-09-856-285-11
; Sequence 11, Application US/09856285
; GENERAL INFORMATION:
; APPLICANT: Ila9, Leodevico L.
; APPLICANT: Ng, Jocelyn, H.
; TITLE OF INVENTION: Method for Modifying and Identifying
; FILE REFERENCE: 50125/024001
; CURRENT APPLICATION NUMBER: US/09/856,285
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/EP99/09052
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: DE 19854196.1
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-856-285-11

```

```

Query Match          100.0%; Score 26; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 FPFY 4
        ||||
Db      3 FPFY 6

```

```

RESULT 11

```

US-07-867-819-86
; Sequence 86, Application US/07867819
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF AUTOANTIBODIES
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,819
; FILING DATE: 13-APR-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 4..11
; US-07-867-819-86

Query Match 100.0%; Score 26; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
|||||
Db 3 PPFY 6

RESULT 12
US-07-867-819A-86
; Sequence 86, Application US/07867819A
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One-Atlantic Center, 1250 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,819A
FILING DATE: 13-APR-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF114CIP(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 4..11
US-07-867-819A-86

Query Match 100.0%; Score 26; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
|||||
Db 3 PPFY 6

RESULT 13
US-07-867-819B-86
; Sequence 86, Application US/07867819B
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: Methods and Reagents for Diagnosis of Autoantibodies
; FILE REFERENCE: OMRF 114 CIP (2)
; CURRENT APPLICATION NUMBER: US/07/867,819B
; CURRENT FILING DATE: 1992-04-13
; PRIOR APPLICATION NUMBER: 07/472,947
; PRIOR FILING DATE: 1990-01-31
; PRIOR APPLICATION NUMBER: 07/648,205
; PRIOR FILING DATE: 1991-01-31
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent version 3.1
; SEQ ID NO 86
; LENGTH: 12
; TYPE: PRT
; ORGANISM: homo sapien
; NAME/KEY: MISC_FEATURE
; LOCATION: (4)..(11)
; OTHER INFORMATION: Binding site
US-07-867-819B-86

Query Match 100.0%; Score 26; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
|||||
Db 3 PPFY 6

RESULT 14
 US-07-867-819D-86
 : Sequence 86, Application US/07867819D
 : GENERAL INFORMATION:
 : APPLICANT: Harley, John
 : TITLE OF INVENTION: Methods and Reagents for Diagnosis of Autoantibodies
 : FILE REFERENCE: OMR# 114 CIP (2)
 : CURRENT APPLICATION NUMBER: US/07/867,819D
 : CURRENT FILING DATE: 1992-04-13
 : PRIOR APPLICATION NUMBER: 07/472,947
 : PRIOR FILING DATE: 1990-01-31
 : PRIOR APPLICATION NUMBER: 07/648,205
 : PRIOR FILING DATE: 1991-01-31
 : NUMBER OF SEQ ID NOS: 161
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 86
 : LENGTH: 12
 : TYPE: PRT
 : ORGANISM: homo sapien
 : FEATURE:
 : NAME/KEY: MISC_FEATURE
 : LOCATION: (4)..(11)
 : OTHER INFORMATION: Binding site
 US-07-867-819D-86

Query Match 100.0%; Score 26; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PPY 4
 ||||
 Db 3 PPY 6

RESULT 15
 US-08-475-955-86
 : Sequence 86, Application US/08475955
 : GENERAL INFORMATION:
 : APPLICANT: Harley, John
 : TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
 : TITLE OF INVENTION: AUTOANTIBODIES
 : NUMBER OF SEQUENCES: 218
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Patrea L. Pabst
 : STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
 : CITY: Atlanta
 : STATE: GA
 : COUNTRY: USA
 : ZIP: 30309-3450
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/475,955
 : FILING DATE: June 7, 1995
 : CLASSIFICATION: 424
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/867,819
 : FILING DATE: April 13, 1992
 : CLASSIFICATION: 424
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/648,205
 : FILING DATE: January 31, 1991
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/472,947
 : FILING DATE: January 31, 1990
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Pabst, Patrea L.
 : REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: OMR#114CIP(2)DIV
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (404)-873-8794
 : TELEFAX: (404)-873-8795
 : INFORMATION FOR SEQ ID NO: 86:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 12 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : FEATURE:
 : NAME/KEY: Binding-site
 : LOCATION: 4..11
 : US-08-475-955-86

Query Match 100.0%; Score 26; DB 8; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PPY 4
 ||||
 Db 3 PPY 6

Search completed: August 20, 2003, 13:13:42
 Job time : 42.9398 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 1.15663 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26

Sequence: 1 PPFV 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCR_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	12	6	US-10-376-121A-86
2	26	100.0	34	6	US-10-273-783-8043
3	26	100.0	59	6	US-10-612-783-4565
4	26	100.0	62	6	US-10-603-113-25825
5	26	100.0	67	6	US-10-603-113-23128
6	26	100.0	68	6	US-10-603-113-27267
7	26	100.0	75	6	US-10-617-320-3303
8	26	100.0	76	6	US-10-603-113-27139
9	26	100.0	77	6	US-10-603-114-8012
10	26	100.0	80	6	US-10-612-783-5205
11	26	100.0	94	6	US-10-631-402-2116
12	26	100.0	94	6	US-10-631-441-2116
13	26	100.0	99	6	US-10-612-783-5529
14	26	100.0	103	6	US-10-617-978-113
15	26	100.0	107	6	US-10-612-783-5576
16	26	100.0	118	6	US-10-612-783-6421
17	26	100.0	125	6	US-10-603-114-7177
18	26	100.0	142	6	US-10-603-113-15390
19	26	100.0	143	6	US-10-612-783-6098
20	26	100.0	146	6	US-10-612-783-6380
21	26	100.0	176	6	US-10-603-113-20869
22	26	100.0	179	6	US-10-612-783-4817
23	26	100.0	184	6	US-10-286-897-1962
24	26	100.0	184	6	US-10-258-898A-1962
25	26	100.0	189	6	US-10-603-113-19356
26	26	100.0	200	6	US-10-286-897-5534

27	26	100.0	200	6	US-10-258-898A-5534	Sequence 5534, Ap
28	26	100.0	201	6	US-10-286-897-2801	Sequence 2801, Ap
29	26	100.0	201	6	US-10-258-898A-2801	Sequence 2801, Ap
30	26	100.0	215	6	US-10-286-897-6373	Sequence 6373, Ap
31	26	100.0	215	6	US-10-258-898A-6373	Sequence 6373, Ap
32	26	100.0	218	6	US-10-617-320-3679	Sequence 3679, Ap
33	26	100.0	225	7	US-60-487-610-2543	Sequence 2543, Ap
34	26	100.0	225	7	US-60-487-610-2544	Sequence 2544, Ap
35	26	100.0	243	6	US-10-603-113-17173	Sequence 17173, A
36	26	100.0	252	6	US-10-408-765A-2772	Sequence 2772, Ap
37	26	100.0	257	6	US-10-286-897-2855	Sequence 2855, Ap
38	26	100.0	257	6	US-10-258-898A-2855	Sequence 2855, Ap
39	26	100.0	271	6	US-10-286-897-6427	Sequence 6427, Ap
40	26	100.0	271	6	US-10-258-898A-6427	Sequence 6427, Ap
41	26	100.0	275	6	US-10-603-113-14129	Sequence 14129, A
42	26	100.0	284	6	US-10-617-320-3009	Sequence 3009, Ap
43	26	100.0	289	6	US-10-603-113-18420	Sequence 18420, A
44	26	100.0	299	6	US-10-603-114-5625	Sequence 5625, Ap
45	26	100.0	314	7	US-60-485-450-1379	Sequence 1379, Ap

ALIGNMENTS

RESULT 1
US-10-376-121A-86
Sequence 86, Application US/10376121A
GENERAL INFORMATION:
APPLICANT: Hartley, John
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
AUTOANTIBODIES
NUMBER OF SEQUENCES: 218
CORRESPONDENCE ADDRESS:
ADDRESS: Patrea L. Pabst
STREET: Suite 2000, 1201 West Peachtree Street, N.E.
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3400
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/376,121A
FILING DATE: 27-Mar-2003
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/867,819
FILING DATE: April 13, 1992
APPLICATION NUMBER: 07/648,205
FILING DATE: January 31, 1991
APPLICATION NUMBER: 07/472,947
FILING DATE: January 31, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMF114CIP(2)DIV(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-817-8473
TELEFAX: (404)-817-8588
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Binding-site
LOCATION: 4..11
SEQUENCE DESCRIPTION: SEQ ID NO: 86:

US-10-376-121A-86

Query Match 100.0%; Score 26; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 5,7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
1111
DB 3 FPFY 6

RESULT 2

US-10-273-573-8043
; Sequence 8043, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8043
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-8043

Query Match 100.0%; Score 26; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
1111
DB 3 FPFY 6

RESULT 3

US-10-612-783-4565
; Sequence 4565, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 4565
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4577_135919C.1.pap
US-10-612-783-4565

Query Match 100.0%; Score 26; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
1111
DB 53 FPFY 56

RESULT 4

US-10-603-113-25825

; Sequence 25825, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 25825
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-25825

Query Match 100.0%; Score 26; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
1111
DB 8 FPFY 11

RESULT 5

US-10-603-113-23128
; Sequence 23128, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 23128
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-23128

Query Match 100.0%; Score 26; DB 6; Length 67;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
1111
DB 46 FPFY 49

RESULT 6

US-10-603-113-27267
; Sequence 27267, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 27267
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-27267

Query Match 100.0%; Score 26; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
1111
DB 23 PPFY 26

RESULT 7

US-10-617-320-3303
Sequence 3303, Application US/10617320
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS

NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/617,320
FILING DATE: 10-Jul-2003

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998

APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011
TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3303:

SEQUENCE CHARACTERISTICS:

LENGTH: 75 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc-feature

LOCATION: (B) LOCATION 1..75

SEQUENCE DESCRIPTION: SEQ ID NO: 3303:

US-10-617-320-3303

Query Match 100.0%; Score 26; DB 6; Length 75;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
1111
DB 39 PPFY 42

RESULT 8

US-10-603-113-27139
Sequence 27139, Application US/10603113
GENERAL INFORMATION:

APPLICANT: Keith Weinstein et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/10/603,113

CURRENT FILING DATE: 2003-06-24

PRIOR APPLICATION NUMBER: US/09/248,796

PRIOR FILING DATE: 1999-02-12

NUMBER OF SEQ ID NOS: 28206

SEQ ID NO 27139

LENGTH: 76

TYPE: PRT

ORGANISM: Candida albicans

US-10-603-113-27139

Query Match 100.0%; Score 26; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
1111
DB 25 PPFY 28

RESULT 9

US-10-603-114-8012
Sequence 8012, Application US/10603114
GENERAL INFORMATION:

APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRA
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1002-001

CURRENT APPLICATION NUMBER: US/10/603,114

CURRENT FILING DATE: 2003-06-24

PRIOR APPLICATION NUMBER: US/09/543,681

PRIOR FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 8012

LENGTH: 77

TYPE: PRT

ORGANISM: Proteus mirabilis

US-10-603-114-8012

Query Match 100.0%; Score 26; DB 6; Length 77;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
1111
DB 31 PPFY 34

RESULT 10

US-10-612-783-5205
Sequence 5205, Application US/10612783
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53373)A

CURRENT APPLICATION NUMBER: US/10/612,783

CURRENT FILING DATE: 2003-07-02

NUMBER OF SEQ ID NOS: 7098

SEQ ID NO 5205

LENGTH: 80

;; TYPE: PRT
;; ORGANISM: zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4577_179465C.1.pep
US-10-612-783-5205

Query Match 100.0%; Score 26; DB 6; Length 80;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
DB 15 FPFY 18

RESULT 11
US-10-631-402-2116
; Sequence 2116, Application US/10631402
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,402
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 2116
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Hypertrophic prostate
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -24...-1
; OTHER INFORMATION: Von Heijne matrix
; FEATURE:
; OTHER INFORMATION: score 6.60
; FEATURE:
; OTHER INFORMATION: seq KMWHLVLGANG/MQ
; NAME/KEY: UNSURE
; LOCATION: 45
; OTHER INFORMATION: Xaa = Phe,Ile,Leu,Val
US-10-631-402-2116

Query Match 100.0%; Score 26; DB 6; Length 94;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
DB 54 FPFY 57

RESULT 12
US-10-631-441-2116
; Sequence 2116, Application US/10631441
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-T119C1
; CURRENT APPLICATION NUMBER: US/10/631,441
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/547,599C
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 2116
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Hypertrophic prostate
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -24...-1
; OTHER INFORMATION: Von Heijne matrix
; FEATURE:
; OTHER INFORMATION: score 6.60
; FEATURE:
; OTHER INFORMATION: seq KMWHLVLGANG/MQ
; NAME/KEY: UNSURE
; LOCATION: 45
; OTHER INFORMATION: Xaa = Phe,Ile,Leu,Val
US-10-631-441-2116

Query Match 100.0%; Score 26; DB 6; Length 94;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
DB 54 FPFY 57

RESULT 13
US-10-612-783-5529
; Sequence 5529, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53373)A

;; CURRENT APPLICATION NUMBER: US/10/612,783
;; CURRENT FILING DATE: 2003-07-02
;; NUMBER OF SEQ ID NOS: 7098
;; SEQ ID NO 5529
;; LENGTH: 99
;; TYPE: PRT
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(99)
;; OTHER INFORMATION: unsure at all Xaa locations
;; FEATURE:
US-10-612-783-5529
OTHER INFORMATION: Clone ID: PAT_MRT4577_36430C.1.pep

Query Match 100.0%; Score 26; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
Db 54 FPFY 57

RESULT 14
US-10-617-978-13
;; Sequence 13, Application US/10617978
;; GENERAL INFORMATION:
;; APPLICANT: Herrmann, Rafael
;; APPLICANT: Lu, Albert L.
;; APPLICANT: McCutchen, Billy Fred
;; APPLICANT: Presnall, James K.
;; APPLICANT: Wong, James F. H.
;; TITLE OF INVENTION: Orally Active Pesticidal Bioloepptides
;; FILE REFERENCE: 35718/260673
;; CURRENT APPLICATION NUMBER: US/10/617,978
;; CURRENT FILING DATE: 2003-07-11
;; PRIOR APPLICATION NUMBER: 60/395,428
;; PRIOR FILING DATE: 2002-07-12
;; NUMBER OF SEQ ID NOS: 27
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 103
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: PRT signal peptide linked to VCI
US-10-617-978-13

Query Match 100.0%; Score 26; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
Db 7 FPFY 10

RESULT 15
US-10-612-783-5576
;; Sequence 5576, Application US/10612783
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;; FILE REFERENCE: 38-21(53373)A
;; CURRENT APPLICATION NUMBER: US/10/612,783
;; CURRENT FILING DATE: 2003-07-02
;; NUMBER OF SEQ ID NOS: 7098
;; SEQ ID NO 5576

;; LENGTH: 107
;; TYPE: PRT
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(107)
;; OTHER INFORMATION: unsure at all Xaa locations
;; FEATURE:
US-10-612-783-5576
OTHER INFORMATION: Clone ID: PAT_MRT4577_39173C.1.pep

Query Match 100.0%; Score 26; DB 6; Length 107;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
Db 60 FPFY 63

Search completed: August 20, 2003, 12:45:14
Job time : 2.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 2.79518 Seconds
(without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26

Sequence: 1 PPFY 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	45	2	EB2117
2	26	100.0	50	2	D90403
3	26	100.0	57	1	A32541
4	26	100.0	68	2	C70246
5	26	100.0	73	2	T16615
6	26	100.0	76	2	S51629
7	26	100.0	77	2	AE3043
8	26	100.0	92	2	E80483
9	26	100.0	97	2	A99427
10	26	100.0	97	2	C56556
11	26	100.0	99	2	B96969
12	26	100.0	101	2	A47450
13	26	100.0	101	2	T60918
14	26	100.0	101	2	T60919
15	26	100.0	105	2	F90335
16	26	100.0	106	2	S51626
17	26	100.0	106	2	S51630
18	26	100.0	111	2	FA7746
19	26	100.0	111	2	E56556
20	26	100.0	116	2	S77926
21	26	100.0	117	2	I49674
22	26	100.0	118	2	C64663
23	26	100.0	118	2	G71851
24	26	100.0	119	2	S69293
25	26	100.0	123	2	A69884
26	26	100.0	128	2	E46178
27	26	100.0	128	2	A50471
28	26	100.0	141	2	S29983
29	26	100.0	142	2	T39842

30	26	100.0	146	2	T26379	hypothetical prote
31	26	100.0	147	2	D32804	chorismate mutase
32	26	100.0	147	2	A11030	probable lipoprote
33	26	100.0	153	2	E71526	hypothetical prote
34	26	100.0	159	1	Q6ADE	early E3 18.5K gly
35	26	100.0	160	1	ERADA5	conserved hypotet
36	26	100.0	161	2	G81544	hypothetical prote
37	26	100.0	162	2	A84217	multidrug resistan
38	26	100.0	164	2	C69842	hypothetical prote
39	26	100.0	164	2	C96841	hypothetical prote
40	26	100.0	164	2	G71427	hypothetical prote
41	26	100.0	166	2	H70056	hypothetical prote
42	26	100.0	166	2	F96024	conserved hypotet
43	26	100.0	174	1	H64007	hypothetical prote
44	26	100.0	175	2	F71982	hypothetical prote
45	26	100.0	176	2	T50522	hypothetical prote

ALIGNMENTS

RESULT 1
EB2117
hypothetical protein VC2102 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: EB2117
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
Chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dregoi, I.; Sellers
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: EB2117
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-45 <HEI>
A:Cross-references: GB:AE004284; GB:AE003852; NID:9656649; PTDN:AAF95248.1; GSPDB:GN
C:Genetics:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Gene: VC2102
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4
DB 26 PPFY 29

RESULT 2
D90403
hypothetical protein SS010228 [imported] - Sulfolobus solfataricus transposon ISC1058
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: D90403
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: D90403
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <KUR>
A:Cross-references: GB:AE006641; NID:q13815629; PTDN:AAK42483.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS010228

Query Match 100.0%; Score 26; DB 2; Length 50;

Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
|||||
DB 25 FPFY 28

RESULT 3

A32541
histatin 1 precursor (validated) - human
N:Alternate names: histidine-rich protein 1
N:Contents: histatin 2
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-2000
R:Accession: I57425; A32541; A22987; A25661; A60664; A28164; A60742; A60659
R:Sabatini, L.M.; Ota, T.; Azen, E.A.
Mol. Biol. Evol. 10, 497-511, 1993
A:Title: Nucleotide sequence analysis of the human salivary protein genes HIS1 and HIS2.
A:Reference number: I57425; MUID:93330039; PMID:8336540
A:Accession: I57425
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-57 <SAL>
A:Cross-references: GB:L04132; NID:g184051; PIDN:AAA02745.1; PID:g184054
R:Sabatini, L.M.; Azen, E.A.
Biochem. Biophys. Res. Commun. 160, 495-502, 1989
A:Title: Histatins, a family of salivary histidine-rich proteins, are encoded by at least
A:Reference number: A90156; MUID:89246491; PMID:2719677
A:Accession: A32541
A:Molecule type: mRNA
A:Residues: 1-57 <SAB>
A:Cross-references: GB:A26664; NID:g292143; PIDN:AAA58645.1; PID:g292144
R:Vanderspek, J.C.; Wyandt, H.E.; Skare, J.C.; Milunsky, A.; Oppenheim, F.G.; Troxler, F.
Am. J. Hum. Genet. 45, 381-387, 1989
A:Title: Localization of the genes for histatins to human chromosome 4q13 and tissue dist
A:Reference number: A32987; MUID:89371745; PMID:2773933
A:Accession: A32987
A:Molecule type: mRNA
A:Residues: 14-57 <YAN>
R:Oppenheim, F.G.; Yang, Y.C.; Diamond, R.D.; Hyslop, D.; Offner, G.D.; Troxler, R.F.
J. Biol. Chem. 261, 1177-1182, 1986
A:Title: The primary structure and functional characterization of the neutral histidine-
A:Reference number: A25661; MUID:86111755; PMID:3944083
A:Accession: A25661
A:Molecule type: protein
A:Residues: 20-57 <OPP>
R:Oppenheim, F.G.; Xu, T.; McMillian, F.M.; Levitz, S.M.; Diamond, R.D.; Offner, G.D.; T
J. Biol. Chem. 263, 7472-7477, 1988
A:Title: Histatins, a novel family of histidine-rich proteins in human parotid secretion
A:Reference number: A94685; MUID:88227937; PMID:3286634
A:Accession: A28164
A:Molecule type: protein
A:Residues: 20-57 <OP2>
R:Troxler, R.F.; Offner, G.D.; Xu, T.; Vanderspek, J.C.; Oppenheim, F.G.
J. Dent. Res. 69, 2-6, 1990
A:Title: Structural relationship between human salivary histatins.
A:Reference number: A60742; MUID:90154430; PMID:2303555
A:Accession: A60742
A:Molecule type: protein
A:Residues: 31-57 <TRO>
R:Vanderspek, J.C.; Offner, G.D.; Troxler, R.F.; Oppenheim, F.G.
Arch. Oral Biol. 35, 137-143, 1990
A:Title: Molecular cloning of human submandibular histatins.
A:Reference number: A60659; MUID:90262442; PMID:2344289
A:Accession: A60659

A:Molecule type: mRNA
A:Residues: 'VWL', 14-57 <VA2>
C:Genetics:

A:Gene: GDB:HTN1
A:Cross-references: GDB:120068; OMIM:142701
A:Map position: 4q13-4q13
A:Introns: 17/3; 24/3; 34/3
C:Superfamily: histatin precursor; statherin/histatin signal sequence homology
C:Keywords: phosphoprotein; saliva
F:1-19/Domain: statherin/histatin signal sequence homology <SHH>
F:20-57/Product: histatin 1 #status experimental <MA1>
F:31-57/Product: histatin 2 #status experimental <MA2>
F:21/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 26; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
|||||
DB 43 FPFY 46

RESULT 4

C70246
hypothetical protein BRJ10 - Lyme disease spirochete plasmid J/1p38
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: C70246
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Fathiga, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quakenbush, J.; Salzberg, S.; Hanson, M.; Vu
B: Bowman, C.; Garland, S.; Fujii, C.; Colton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: C70246
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-68 <KLE>
A:Cross-references: GB:AE000787; NID:g2690175; PIDN:AAAC66097.1; PID:g2690189; TIGR:BB
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 100.0%; Score 26; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
|||||
DB 45 FPFY 48

RESULT 5

T16615
hypothetical protein K10B3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16615
R:Gallung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid K10B3.
A:Reference number: Z18546
A:Accession: T16615
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-73 <GAT>
A:Cross-references: EMBL:U49941; NID:g1206038; PID:AA53870.1; GSPDB:C
A:Experimental source: strain Bristol N2; clone K10B3
C:Genetics:
A:Gene: CESP:K10B3.1

A:Map position: X
A:Introns: 25/3

Query Match 100.0%; Score 26; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
|||||
DB 52 PPFY 55

RESULT 6

S51629

FREPAC-6 protein - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
C:Accession: S51629
R:Pierrou, S.; Heliqvist, M.; Samuelsson, L.; Enerbaeck, S.; Carlsson, P.
EMBO J. 13, 5002-5012, 1994

A:Title: Cloning and characterization of seven human forkhead proteins: binding site spe

A:Reference number: S51624; MUID:95045392; PMID:7957066

A:Accession: S51629

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-76 <PIE>

A:Cross-references: EMBL:U13224; NID:95563167; PIDN:AAA92041.1; PID:95563168

C:Genetics:

A:Gene: GDB:FKHL10; FREPAC

A:Cross-references: GDB:450228; OMIM:601093

A:Map position: 5q34-5q34

C:Superfamily: unassigned fork head proteins: fork head DNA-binding domain homology (FHD)

F:6-76/Domain: fork head DNA-binding domain homology (fragment) <FHD>

Query Match 100.0%; Score 26; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
|||||
DB 40 PPFY 43

RESULT 7

AE2043

hypothetical protein asr1899 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AE2043

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2043

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW3598.1; PID:g17130989; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

Query Match 100.0%; Score 26; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
|||||
DB 19 PPFY 22

RESULT 8

E90483

first ORF in transposase ISCI058 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: E90483

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: E90483

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-92 <KUR>

A:Cross-references: GB:AE006641; NID:913816417; PIDN:AAK43124.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS011734

Query Match 100.0%; Score 26; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
|||||
DB 25 PPFY 28

RESULT 9

A99427

partial transposase ISCI058 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: A99427

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: A99427

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <KUR>

A:Cross-references: GB:AE006641; NID:913815845; PIDN:AAK42672.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS010604

Query Match 100.0%; Score 26; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
|||||
DB 25 PPFY 28

RESULT 10

C56556

fork head homolog XFD-2' - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 31-Oct-1997

C:Accession: C56556

R:Knöchel, S.; Lef, J.; Clement, J.; Klocke, B.; Hille, S.; Koster, M.; Knöchel, W.

Mech. Dev. 38, 157-165, 1992

A:Title: Activin A induced expression of a fork head related gene in posterior chorda

A:Reference number: A56556; MUID:93041288; PMID:1358174

A:Accession: C56556

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-97 <KNO>

A:Experimental source: gastrula

A:Note: sequence extracted from NCBI backbone (NCBIP:118179)
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 97;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
DB 44 FPFY 47

RESULT 11
B96969

Probable membrane protein [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: B96969

R:Cloning, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A:Reference number: A96900; PMID:21359325; PMID:21359325

A:Accession: B96969

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-99 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78541.1; PID:g15023429; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0562

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
DB 30 FPFY 33

RESULT 12
A47450

HNF-3/fork-head homolog-3 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 10-May-1996 #sequence_revision 10-May-1996 #text_change 12-Sep-1997

C:Accession: A47450

R:Cloning, D.E.; Overdier, D.G.; Tao, W.; Qian, X.; Pan, L.; Lai, E.; Costa, R.H.

Proc. Natl. Acad. Sci. U.S.A. 90, 3948-3952, 1993

A:Title: Identification of nine tissue-specific transcription factors of the hepatocyte

A:Reference number: A47450; PMID:93248207; PMID:7683413

A:Accession: A47450

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-101 <RES>

A:Cross-references: GB:U13203; NID:g306843; PID:g306844

C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

C:Keywords: DNA binding; transcription factor

F:3-94/Domain: fork head DNA-binding domain homology <FHD>

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 101;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
DB 37 FPFY 40

RESULT 13
I60918

brain factor-3 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I60918

R:Cloning, D.E.; Overdier, D.G.; Tao, W.; Qian, X.; Pan, L.; Lai, E.; Costa, R.H.

Proc. Natl. Acad. Sci. U.S.A. 90, 3948-3952, 1993

A:Title: Identification of nine tissue-specific transcription factors of the hepatocyte

A:Reference number: A47450; PMID:93248207; PMID:7683413

A:Accession: I60918

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-101 <RES>

A:Cross-references: GB:U13193; NID:g310156; PIDN:AAA41320.1; PID:g310157

C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

F:3-94/Domain: fork head DNA-binding domain homology <FHD>

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 101;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
DB 37 FPFY 40

RESULT 14
I60919

HNF-3/fork-head homolog-5 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I60919

R:Cloning, D.E.; Overdier, D.G.; Tao, W.; Qian, X.; Pan, L.; Lai, E.; Costa, R.H.

Proc. Natl. Acad. Sci. U.S.A. 90, 3948-3952, 1993

A:Title: Identification of nine tissue-specific transcription factors of the hepatocyte

A:Reference number: A47450; PMID:93248207; PMID:7683413

A:Accession: I60919

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-101 <RES>

A:Cross-references: GB:U13205; NID:g310158; PIDN:AAA41321.1; PID:g310159

C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

F:3-94/Domain: fork head DNA-binding domain homology <FHD>

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 101;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
DB 37 FPFY 40

RESULT 15
F90335

hypothetical protein SS01745 [imported] - Sulfolobus solfataricus transposon ISCI058

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: F90335

R:Sheng, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awghey, M.J.; Ch

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to Genbank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: F90335

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <KUR>

A:Cross-references: GB:AE006641; NID:g13814989; PIDN:AAK41941.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS01745

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 105;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	1 FPPY 4
Db	25 FPPY 28

Search completed: August 20, 2003, 12:42:23
Job time : 5.79518 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 1.51807 Seconds
(without alignments) 123.912 Million cell updates/sec

Title: US-09-512-082-31

Perfect score: 26

Sequence: 1 FPPY 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	57	HIS1_HUMAN	P15515 homo sapien
2	26	100.0	90	FXL2_MOUSE	O88470 mus musculu
3	26	100.0	101	FXC2_RAT	O63246 rattus norv
4	26	100.0	101	FXI1_RAT	O63248 rattus norv
5	26	100.0	118	RLI9_HELP1	O92K67 heliobacte
6	26	100.0	118	RLI9_HELP2	P56044 heliobacte
7	26	100.0	138	PA23_ECHOC	P59171 echis ocelli
8	26	100.0	147	PHEB_BACSV	P21204 bacillus su
9	26	100.0	159	E3GL_ADE02	P03251 human adeno
10	26	100.0	160	E3GL_ADE05	P04494 human adeno
11	26	100.0	174	V453_HAEIN	P43999 haemophilus
12	26	100.0	185	DPI_HUMAN	O00765 homo sapien
13	26	100.0	216	RLI9_HELP1	P43332 drosophila
14	26	100.0	225	RLI9_HELP2	P08579 homo sapien
15	26	100.0	226	PRR2_TOBAC	P07052 nicotiana t
16	26	100.0	229	YSV4_CAEEL	O10010 caenorhabdi
17	26	100.0	237	L131_CAEEL	O03256 caenorhabdi
18	26	100.0	240	SRY_CAPHI	O03257 ovis aries
19	26	100.0	240	SRY_SHEEP	O57827 methanococc
20	26	100.0	245	Y382_METUA	O59188 borrelia bu
21	26	100.0	258	PLSC_BORBU	O67773 aquifex aeo
22	26	100.0	271	FLIR_AQUAE	P32029 drosophila
23	26	100.0	271	FDS_DROME	O66861 aquifex aeo
24	26	100.0	272	V603_AQUAE	P20645 homo sapien
25	26	100.0	277	MRPD_HUMAN	P09012 homo sapien
26	26	100.0	282	RLI9_HELP1	P45429 xenopus lae
27	26	100.0	282	RLI9_HELP2	O94092 yaba monkey
28	26	100.0	287	MCE5_YABAM	O21268 caenorhabdi
29	26	100.0	290	YMSO_CAEEL	P56627 rattus norv
30	26	100.0	295	AOP9_RAT	O63623 dalbuis ch
31	26	100.0	310	NTJM_DALCH	P47435 mycoplasma
32	26	100.0	318	Y189_MYCGE	P75262 mycoplasma
33	26	100.0	319	Y189_MYCPN	

34	26	100.0	322	1	RIR2_PLAAG	P50649 plasmodium
35	26	100.0	329	1	ID12_STRPY	O9A095 streptococc
36	26	100.0	336	1	ID12_STRPN	O97A98 streptococc
37	26	100.0	337	1	FXL1_MOUSE	O64731 mus musculu
38	26	100.0	345	1	FXL1_HUMAN	O12952 homo sapien
39	26	100.0	347	1	ID12_LACLA	O9C1F5 lactococcus
40	26	100.0	349	1	SGFI_BOMO	O17241 bombyx mori
41	26	100.0	351	1	FXI1_HUMAN	O12951 homo sapien
42	26	100.0	353	1	EDG7_HUMAN	O94095 homo sapien
43	26	100.0	354	1	EDG7_MOUSE	O94091 mus musculu
44	26	100.0	354	1	EDG7_RAT	O8K360 rattus norv
45	26	100.0	358	1	ID12_LISTIN	O92BX2 listeria ln

ALIGNMENTS

RESULT 1
HIS1_HUMAN STANDARD: PRT; 57 AA.
ID HIS1_HUMAN
AC P15515;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Histatin 1 precursor (Histidine-rich protein 1) (Post-PB protein)
DE (PDB) [Contains: Histatin 2].
GN HTM1 OR HIS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE-69246491; PubMed-2719677;
EX Sabatini L.M., Azen E.A.;
RT "Histatins, a family of salivary histidine-rich proteins, are encoded
RT by at least two loci (HIS1 and HIS2).";
RL Biochem. Biophys. Res. Commun. 160:495-502(1989).
RN [2]
RP MEDLINE-93330039; PubMed-8336540;
EX Chen Z.W.;
RT "Nucleotide sequence analysis of the human salivary protein genes
RT HIS1 and HIS2, and evolution of the STATH/HIS gene family.";
RL Mol. Biol. Evol. 10:497-511(1993).
RN [3]
RP MEDLINE-12477932; PubMed-12477932;
EX Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshimune S., Carrincci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP [4]
RP MEDLINE-86227937; PubMed-3286634;
RC TISSUE-Parotid gland;
RX MEDLINE-86227937; PubMed-3286634;

RA Oppenheim F.G., Xu T., McMillian F.M., Levitz S.M., Diamond R.D.,
 RA Offner G.D., Troxler R.F.;
 RT "Histatins, a novel family of histidine-rich proteins in human
 RT parotid secretion. Isolation, characterization, primary structure,
 RT and fungistatic effects on *Candida albicans*.";
 RL J. Biol. Chem. 263:7472-7477(1988).
 RN [5]
 RP SEQUENCE OF 20-57.
 RC TISSUE-Parotid gland;
 RX MEDLINE=86111755; PubMed=3944083;
 RA Oppenheim F.G., Yang Y.C., Diamond R.D., Hyslop D., Offner G.D.,
 RA Troxler R.F.;
 RT "The primary structure and functional characterization of the neutral
 RT histidine-rich polypeptide from human parotid secretion.";
 RL J. Biol. Chem. 261:1177-1182(1986).
 RN [6]
 RP SEQUENCE OF 20-57.
 RC TISSUE-Saliva;
 RX MEDLINE=90321151; PubMed=2372245;
 RA Sugiyama K., Ogino T., Ogata K.;
 RT "Rapid purification and characterization of histatins (histidine-rich
 RT polypeptides) from human whole saliva.";
 RL Arch. Oral Biol. 35:415-419(1990).
 RN [7]
 RP SEQUENCE OF 14-57 FROM N.A.
 RX MEDLINE=89371745; PubMed=2773933;
 RA Vanderspek J.C., Wyandt H.E., Skare J.C., Milunsky A., Oppenheim F.G.,
 RA Troxler R.F.;
 RT "Localization of the genes for histatins to human chromosome 4q13 and
 RT tissue distribution of the mRNAs.";
 RL Am. J. Hum. Genet. 45:381-387(1989).
 CC -1- FUNCTION: HISTATINS ARE SALIVARY PROTEINS THAT ARE CONSIDERED TO
 CC BE MAJOR PRECURSORS OF THE PROTECTIVE PROTEINACHOUS STRUCTURE ON
 CC TOOTH SURFACES (ENAMEL PELLICLE). IN ADDITION, HISTATINS EXHIBIT
 CC ANTI-BACTERIAL AND ANTI-FUNGAL ACTIVITIES.
 CC -1- PTM: HISTATIN 2 IS A PROTEOLYTIC PRODUCT OF HISTATIN 1.
 CC -1- SIMILARITY: TO HISTATIN 3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M2664; AAA58645.1; -
 DR EMBL: L04132; AAA02745.1; -
 DR EMBL: BC017835; AAH17835.1; -
 DR PIR: I57425; A32541.
 DR Genew; HGNC:5283; HTNL.
 DR MIM: 142701; -
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003797; F:antibacterial peptide activity; NAS.
 DR GO: GO:0006952; P:defense response; NAS.
 KW Bionmineralization; Fungicide; Antibiotic; Signal; Phosphorylation.
 FT SIGNAL 1 19
 FT PEPTIDE 20 57 HISTATIN 1.
 FT MOD_RES 31 57 HISTATIN 2.
 FT MOD_RES 21 21 PHOSPHORYLATION.
 SQ SEQUENCE 57 AA: 6963 MW: 73532BDIDCE23D83 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
 DB 43 PPFY 46

RESULT 2
 FXL2 MOUSE

ID FXL2 MOUSE STANDARD; PRT; 90 AA.
 AC 088470;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Forkhead box protein L2 (Pituitary forkhead factor) (P-Frk)
 DE (Fragment).
 GN FOXL2 OR PFRK.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98283916; PubMed=9620855;
 RA Treier M., Gleiberman A.S., O'Connell S.M., Szeto D.P., McMahon J.A.,
 RA McMahon A.P., Rosenfield M.G.;
 RT "Multistep signaling requirements for pituitary organogenesis in
 RT vivo.";
 RL Genes Dev. 12:1691-1704(1998).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT OVARY AND IN THE
 CC MESENCHYME OF DEVELOPING EYELIDS.
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF060873; AAC27508.1; -
 DR HSSP: Q63245; 2HFH.
 DR MGD: MGI:1349428; Foxl2.
 DR InterPro: IPR001766; TF_Fork_head.
 DR Pfam: PF00250; Fork_head_1.
 DR PRINTS: PR00053; FORKHEAD.
 DR PRODOM: PD000425; TF_Fork_head; 1.
 DR SMART: SM00339; FH; 1.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; 1.
 DR PROSITE: PS0039; FORK_HEAD_3; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation.
 FT NON_TER 1 1
 FT DNA_BIND 1 >90 FORK-HEAD.
 FT NON_TER 90 90
 SQ SEQUENCE 90 AA: 10484 MW: 03CE0DEC6BA44EBB CRC64;

Query Match 100.0%; Score 26; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPFY 4
 DB 35 PPFY 38

RESULT 3
 FXC2-RAT STANDARD; PRT; 101 AA.
 ID FXC2-RAT
 AC 063246;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Forkhead box protein C2 (Brain factor-3) (BF-3) (HFH-BF-3) (Fragment).
 GN FOXC2 OR HFHBF3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; Tao W., Chen W.S., Darnell J.E. Jr.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MIGHT BE INVOLVED IN THE FORMATION OF SPECIAL
CC MESENCHYMAL TISSUES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: L13193; AAA41320.1; -
DR PIR: I60918; I60918.
DR HSSP: O63245; 2HRH.
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR ProDom: PD000425; TF_Fork_head; 1.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
DR Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 1 2 93 FORK-HEAD.
FT NON_TER 1 101
FT SEQUENCE 101 AA: 12098 MW: 78BED5BEF53F801 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 101;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
DB 37 FFY 40

RESULT 4
EXIL_RAT
ID EXIL_RAT STANDARD: PRT; 101 AA.
AC 063248;
DR 01-NOV-1997 (Rel. 35, Created)
DR 01-NOV-1997 (Rel. 35, Last sequence update)
DR 30-MAY-2000 (Rel. 39, Last annotation update)
DE Forkhead box protein 11 (Forkhead-related protein FKHL10) (Forkhead-
DE related transcription factor 5) (FH-5) (Fragment).
DE Factor 3 forkhead homolog 5 (FH-5) (Fragment).
GN FOX11 OR FKHL10 OR FREAC6 OR HPH5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;

RN
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley;
RC MEDLINE-93248207; PubMed-7683413;
RA Cleveland D.E., Overdier D.G., Tao W., Qian X., Paul L., Lal E.,
RA Costa R.H.;
RT Identification of nine tissue-specific transcription factors of the
RT hepatocyte nuclear factor 3/forkhead DNA-binding-domain family.;
RL Proc. Natl. Acad. Sci. U.S.A. 90:3948-3952(1993).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: L13205; AAA41321.1; -
DR PIR: I60919; I60919.
DR HSSP: O63245; 2HRH.
DR TRANSFAC: T02480; -
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR ProDom: PD000425; TF_Fork_head; 1.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
DR DNA-binding; Nuclear protein; Transcription regulation; Activator.
FT DNA_BIND 1 2 93 FORK-HEAD.
FT NON_TER 1 101
FT SEQUENCE 101 AA: 11988 MW: 3BED48CF13A26F9C CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 101;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
DB 37 FFY 40

RESULT 5
RL19_HELPJ
ID RL19_HELPJ STANDARD: PRT; 118 AA.
AC Q92K67;
DR 30-MAY-2000 (Rel. 39, Created)
DR 30-MAY-2000 (Rel. 39, Last sequence update)
DR 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L19.
GN RPLS OR JRP1074.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_Taxid=85963;

RN
RP SEQUENCE FROM N.A.
RC MEDLINE-99120557; PubMed-9923682;
RA Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Gull B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Ulla-Mckelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust J.J.;
RT Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.;
RL Nature 397:176-180(1999).
CC -1- FUNCTION: THIS PROTEIN IS LOCATED AT THE 30S-50S RIBOSOMAL SUBUNIT
CC INTERFACE AND MAY PLAY A ROLE IN THE STRUCTURE AND FUNCTION OF THE
CC AMINOACYL-TRNA BINDING SITE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L19 FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: AE001535; AAD06654.1; -
DR PIR: G71851; G71851.
DR HAMAP: MF_00402; -; 1.

```

DR InterPro: IPR001857; Ribosomal_L19.
 DR Pfam: PF01245; Ribosomal_L19; 1.
 DR PRINTS: PR00061; RIBOSOMAL_L19.
 DR ProDom: PD002979; Ribosomal_L19; 1.
 DR TIGRFAMs: TIGR01024; rplS_bact; 1.
 DR PROSITE: PS01015; RIBOSOMAL_L19; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 118 AA; 13616 MW; 8B21B63CAD45AC5 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 118;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
 ||||
 Db 77 FFY 80

RESULT 6

RL19_HELPY STANDARD; PRT; 118 AA.
 ID RL19_HELPY
 AC P56044;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L19.
 GN RPL19 OR HPL147.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McEnney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback J.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 pylori.";
 RL Nature 388:539-547(1997).

-1- FUNCTION: THIS PROTEIN IS LOCATED AT THE 30S-50S RIBOSOMAL SUBUNIT
 INTERFACE AND MAY PLAY A ROLE IN THE STRUCTURE AND FUNCTION OF THE
 AMINOACYL-TRNA BINDING SITE (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE L19P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE000621; AAD08191.1; -
 DR PIR: C64663; C64663.
 DR TIGR: HPL147; -
 DR HAMAP: MF_00402; -; 1.
 DR InterPro: IPR001857; Ribosomal_L19.
 DR Pfam: PF01245; Ribosomal_L19; 1.
 DR PRINTS: PR00061; RIBOSOMAL_L19.
 DR ProDom: PD002979; Ribosomal_L19; 1.
 DR TIGRFAMs: TIGR01024; rplS_bact; 1.
 DR PROSITE: PS01015; RIBOSOMAL_L19; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 118 AA; 13600 MW; A151AD3A165A233F CRC64;

Query Match 100.0%; Score 26; DB 1; Length 118;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
 ||||
 Db 77 FFY 80

RESULT 7

PA25_ECHOC STANDARD; PRT; 138 AA.
 ID PA25_ECHOC
 AC P59171;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phospholipase A2 5 precursor (EC 3.1.1.4) (Phosphatidylcholine 2-
 DE acylhydrolase).
 GN PLA2-5.
 OS Echis ocellatus (Ocellated saw-scaled viper).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Viperidae; Viperinae; Echis.
 OX NCBI_TaxID=99586;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom gland;
 RA Harrison R.A., Hasson S., Bharati K.;
 RT "Molecular cloning of phospholipase A2 from the venom glands of Echis
 RT carpet vipers.";
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
 CC acyl groups in 3-sn-phosphoglycerides (By similarity).
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O -> 1-
 CC acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY. GROUP II
 CC SUBFAMILY.

-1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF539921; AAN77204.1; -
 DR ProDom: PD000303; PhospholipaseA2; 1.
 DR SMART: SM00085; PA2C; 1.
 DR PROSITE: PS00118; PA2_HIS; 1.
 DR PROSITE: PS00119; PA2_ASP; 1.
 KW Hydrolyase; Lipid degradation; Calcium; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 138
 FT ACT_SITE 63 63
 FT ACT_SITE 105 105
 FT DISULFID 42 131
 FT DISULFID 44 60
 FT DISULFID 59 111
 FT DISULFID 65 138
 FT DISULFID 66 104
 FT DISULFID 73 97
 FT DISULFID 91 102
 FT METAL 43 43
 FT METAL 45 45
 FT METAL 47 47
 FT METAL 64 64

SO SEQUENCE 138 AA; 15705 MW; 0896F3507C411AE1 CRC64;
 Query Match 100.0%; Score 26; DB 1; Length 138;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PPEY 4
 Db 34 PPEY 37

RESULT 8
 PHEB_BACSU STANDARD; PRT; 147 AA.
 AC P21204;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Possible chorismate mutase (EC 5.4.99.5) (CM).
 GN PHEB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89153435; PubMed=2537815;
 RA Trich K., Hoch J.A.;
 RT "The Bacillus subtilis spo0B stage 0 sporulation operon encodes an essential GTP-binding protein.";
 RL J. Bacteriol. 171:1362-1371(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Allout G.,
 RA Azevedo V., Bortero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Bortiss R., Bourlier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertlan K.D., Errington J., Fabel C., Ferrari E., Foulger D.,
 RA Fujita Y., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Hilbert H., Holsappel S., Hosono S., Hulls B.F., Itaya M., Jones L.,
 RA Joris B., Karmata D., Kasahara Y., Kleerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 RA Preegan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche M., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serrito P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoli E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpsira P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenol M., Vanlier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Wetzinger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis".
 RT Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: Chorismate - prephenate.
 CC -1- PATHWAY: L-phenylalanine biosynthesis.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M24537; AAA22506.1; -;
 CC EMBL: Z99118; CAB14751.1; -;
 CC PIR: D32804; D32804.
 CC Subtilist: BG10913; pheB.
 CC InterPro: IPR002912; ACT.
 CC Pfam: PF01842; ACT; 1.
 CC Phenylalanine biosynthesis; Isomerase; Complete proteome.
 CC SW Phenylalanine biosynthesis; Isomerase; Complete proteome.
 CC SEQUENCE 147 AA; 16659 MW; 4D11D222BA520E27 CRC64;
 Oy 1 PPEY 4
 Db 58 PPEY 61

RESULT 9
 E3GL_ADE02 STANDARD; PRT; 159 AA.
 AC P03251;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Early E3 18.5 kDa glycoprotein precursor (GP19K).
 OS Human adenovirus type 2, and
 OS Human adenovirus type 6.
 OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_Taxid=10515, 10534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human adenovirus type 2;
 RX MEDLINE=81053687; PubMed=5253880;
 RA Herisse J., Courtois G., Galibert F.;
 RT "Nucleotide sequence of the E3 D fragment of adenovirus 2 genome.";
 RL Nucleic Acids Res. 8:2173-2192(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human adenovirus type 6;
 RA Reichtmann H., Schaarschmidt E., Gelsler B., Hausmann J., Ottmann D.,
 RA Bauer U., Flunker G., Seidel W.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: FORMS A COMPLEX WITH THE HEAVY CHAIN OF THE CLASS I
 CC TRANSLANTATION ANTIGENS DURING AN ADENOVIRUS INFECTION, THEREBY
 CC REDUCING CLASS I EXPRESSION ON THE CELL SURFACE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
 CC reticulum.
 CC -1- SIMILARITY: TO OTHER ADENOVIRUSES GP19K.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entitles requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J01917; AAA92221.1; -;
 CC EMBL: Y16037; CAA75990.1; -;
 CC PIR: A03821; O6ADE.
 CC InterPro: IPR006965; Adeno-GP19K.
 CC Pfam: PF04881; Adeno-GP19K; 1.
 CC Early protein. Transmembrane; Glycoprotein; Signal.
 CC STGNAL. 1 17
 CC CHAIN 18 159
 CC DOMAIN 18 123
 CC TRANSMEM 124 144
 CC DOMAIN 145 159
 CC CARBOHYD 29 29
 CC POTENTIAL. 1
 CC EARLY E3 18.5 KDA GLYCOPROTEIN.
 CC LUMENAL (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 159 AA; 18438 MW; ED2519547E18AE9D CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 26; DB 1; Length 159;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 PPFY 4
 1111
 94 PPFY 97
 Db
 RESULT 10
 E3GL_ADE05 STANDARD; PRT; 160 AA.
 AC P04494;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Early E3 18.5 kDa glycoprotein precursor (GP19K).
 OS Human adenovirus type 5.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 NCBI_TaxID=28285;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85092386; PubMed=2981456;
 RA Cladaras C., Wold W.S.M.;
 RT "DNA sequence of the early E3 transcription unit of adenovirus 5.";
 RL Virology 140:28-43(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85130985; PubMed=3882694;
 RA Wold W.S.M., Cladaras C., Deutscher S.L., Kapoor Q.S.;
 RT "The 19-kDa glycoprotein coded by region E3 of adenovirus.
 Purification, characterization, and structural analysis.";
 RL J. Biol. Chem. 260:2424-2431(1985).
 RN [3]
 RP COMPLETE GENOME.
 RX MEDLINE=92087470; PubMed=1727603;
 RA Chroboczek J., Bleher F., Jacrot B.;
 RT "The sequence of the genome of adenovirus type 5 and its comparison
 with the genome of adenovirus type 2.";
 RL Virology 186:280-285(1992).
 CC -1- FUNCTION: FORMS A COMPLEX WITH THE HEAVY CHAIN OF THE CLASS I
 TRANSLANTION ANTIGENS DURING AN ADENOVIRUS INFECTION, THEREBY
 REDUCING CLASS I EXPRESSION ON THE CELL SURFACE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
 reticulum.
 CC -1- SIMILARITY: TO OTHER ADENOVIRUSES GP19K.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M33260; -; NOT_ANNOTATED_CDS.
 DR EMBL; X03002; CAA26783.1; -.
 DR EMBL; M12406; AAA42492.1; -.
 DR PIR; A03822; ERA0A5.
 DR InterPro: IPR006965; Adeno_GP19K.
 DR Pfam: PF04881; Adeno_GP19K; 1.
 KW Early protein; Transmembrane; Glycoprotein; Signal.
 FT CHAIN 1 18
 FT SIGNAL 1 18
 FT STAIN 1 18
 FT CHAIN 1 160
 FT DOMAIN 19 160
 FT DOMAIN 125 145
 FT TRANSMEM 125 145
 FT DOMAIN 146 160
 FT CARBOHYD 30 30
 FT CARBOHYD 79 79
 FT CARBOHYD 160 AA; 18502 MW; B1P0D2AC4D6330E0 CRC64;
 SQ SEQUENCE

Query Match
 Best Local Similarity 100.0%; Score 26; DB 1; Length 160;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 PPFY 4
 1111
 95 PPFY 98
 Db
 RESULT 11
 Y453_HAEIN STANDARD; PRT; 174 AA.
 ID Y453_HAEIN
 AC P43939;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H10453.
 GN H10453.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 NCBI_TaxID=727;
 [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 Gnehm C.L., McDonald L.A., Small R.V., Fraser C.M., Smith H.O.,
 Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RT Science 269:496-512(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32728; AAC22111.1; -.
 DR PIR; H64007; H64007.
 DR TIGR; H10453; -.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 8 28
 FT TRANSMEM 146 166
 FT SEQUENCE 174 AA; 20416 MW; 32A2F39303399B48 CRC64;
 SQ SEQUENCE
 Query Match
 Best Local Similarity 100.0%; Score 26; DB 1; Length 174;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 PPFY 4
 1111
 87 PPFY 90
 Db
 RESULT 12
 DPL_HUMAN STANDARD; PRT; 185 AA.
 ID DPL_HUMAN
 AC Q00765; Q04198; Q9BWH9;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Polypsis locus protein 1 (Tb2 protein).

GN DPL OR TB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91330307; PubMed-1678319;
 RA Joslyn G., Carlson M., Thilveris A., Albertsen H., Gelbert L.,
 RA Samowitz W., Groden J., Stevens J., Spirio L., Robertson M.,
 RA Sargent L., Krapcho K., Wolff E., Butt R., Hughes J.P.,
 RA Warrington J., McPherson J.D., Wasmuth J.J., Le Paslier D.,
 RA Abderhaldin H., Cohen D., Leppert M., White R.;
 RT "Identification of deletion mutations and three new genes at the
 RT familial polyposis locus.";
 RL Cell 66:601-613(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91335210; PubMed-1651562;
 RA Kinler K.W., Nibert M.C., Su L.K., Vogelstein B., Bryan T.M.,
 RA Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McKechnie D.,
 RA Fimlinger R., Markham A., Groffen J., Boguski M.S., Altschul S.F.,
 RA Horii A.K., Ando H., Miyoshi Y., Miki Y., Nishisho I., Nakamura Y.;
 RT "Identification of FAP locus genes from chromosome 5q21.";
 RL Science 253:661-663(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-EYE;
 RX MEDLINE-22388257; PubMed-12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dichtenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tomshylyk S., Carninci P., Prange C.,
 RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schenker A., Schell J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO C.ELEGANS T1903.4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M73547; AAA60136.1; -
 DR EMBL; M74090; AAA66351.1; ALT-INIT.
 DR EMBL; BC000232; AAH00232.1; -
 DR PIR; A39658; A39658.
 DR MTM; 125265; -
 DR GO; GO:0016021; C:Integral to membrane; NAS.
 DR InterPro; IPR004345; TB2_DPL_HVA22.
 DR Pfam; PF03134; TB2_DPL_HVA22; 1.
 KW Transmembrane.
 FT TRANSMEM 31 51
 FT TRANSMEM 86 106
 FT TRANSMEM 111 111
 FT TRANSMEM 159 159
 FT POTENTIAL.
 FT M -> I (IN REF. 2).
 FT A -> S (IN REF. 1).
 FT CONFLICT 159 159

SQL SEQUENCE 185 AA; 21132 MW; E36B961DA56D2BA0 CRC64;
 Query Match 100.0%; Score 26; DB 1; Length 185;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPFY 4
 DB 106 PPFY 109
 RESULT 13
 ID RUIA_DROME STANDARD; PRT; 216 AA.
 AC P43332; Q9W4D7;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE U1 small nuclear ribonucleoprotein A (U1 snRNP A protein) (Sex
 DE determination protein SNF).
 GN SNF OR D25 OR LIZ OR FS(1)1621 OR CG4528.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92350664; PubMed-1386424;
 RA Harper D.S., Fresco L.D., Keene J.D.;
 RT "RNA binding specificity of a Drosophila snRNP protein that shares
 RT nucleic acid Res. 20:3645-3650(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-Oregon-R;
 RX MEDLINE-95011590; PubMed-7926776;
 RA Flickinger T.W., Salz H.K.;
 RT "The Drosophila sex determination gene snf encodes a nuclear protein
 RT with sequence and functional similarity to the mammalian U1A snRNP
 RT protein.";
 RL Genes Dev. 8:914-925(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN-Berkely;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agayari A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brodeur P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cusum S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., De J., Dietz S.M.,
 RA Dodson K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Idegami C.,
 RA Jajani M., Kalush F., Katpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko F., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtka R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE-22426066; PubMed-12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A *Drosophila* full-length cDNA resource";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.6(2002).
CC -1- FUNCTION: BINDS STEM LOOP II OF U1 snRNA. IT IS THE FIRST SN-RNP
CC TO INTERACT WITH PRE-MRNA. THIS INTERACTION IS REQUIRED FOR THE
CC SUBSEQUENT BINDING OF U2 SN-RNP AND THE U4/U6/U5 TRI-SN-RNP (BY
CC SIMILARITY). PLAYS A ROLE IN REGULATING SEX-LETHAL SPLICING.
CC -1- SUBUNIT: BELONGS TO THE SPLICOSOME WHERE IT IS ASSOCIATED WITH
CC SN-RNP U1
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE U1 A/B FAMILY.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: M89775; AAA28441.1; -;
DR EMBL: L29521; AAA28903.1; -;
DR EMBL: AE003433; AAF46017.1; -;
DR EMBL: AY061491; AAL29039.1; -;
DR PIR: A54279; A54279.
DR HSSP: P09012; 3UTR.
DR FlyBase: FBgn0003449; snf.
DR GO: GO:0030532; C:small nuclear ribonucleoprotein complex; IDA.
DR GO: GO:0007539; P:primary sex determination; soma; NAS.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS50102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP.1; 1.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat; Spliceosome.
FT DOMAIN 7 86
FT DOMAIN 142 216
FT VARIANT 49 49 R -> H (IN ALLELE SNF121; STERILE).
SQ SEQUENCE 216 AA; 24546 MW; 5B736FEF836523373 CRC64;
Query Match 100.0%; Score 26; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FPFY 4
DB 72 FPFY 75
RESULT 14
RUB2 HUMAN STANDARD; PRT; 225 AA.
AC P08579; Q9UDJ4;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE U2 small nuclear ribonucleoprotein B".
GN SNRBP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-87175685; PubMed-2951739;
RA Habets W.J., Sililekens P.T.G., Hoet M.H., Schaiken J.A.,
RA Roebroek A.J.M., Leunissen J.A.M., de Ven W.J.M., van Verrooij W.J.;
RT "Analysis of a cDNA clone expressing a human autoantigen:
RT full-length sequence of the U2 small nuclear RNA-associated B'
RT antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2421-2425(1987).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-21638749; PubMed-11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharasialho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachle L.J., McLeay K., Murray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Snowken R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Whiting L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RN IDENTIFICATION IN SPLICOSOMAL COMPLEX WITH HPR8BP AND CRNL1.
RX MEDLINE-22079017; PubMed-12084575;
RA Chung S., Zhou Z., Huddleston K.A., Harrison D.A., Reed R.,
RA Coleman T.A., Rymond B.C.;
RT "Crooked neck is a component of the human spliceosome and implicated
RT in the splicing process.";
RL Biochim. Biophys. Acta 1576:287-297(2002).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (2.38 ANGSTROMS) OF 1-96.
RX MEDLINE-98379985; PubMed-9716128;
RA Price S.R., Evans P.R., Nagai K.;
RT "Crystal structure of the spliceosomal U2B'-U2A' protein complex
RT bound to a fragment of the small nuclear RNA.";
RL Nature 394:645-650(1998).
CC -1- FUNCTION: Involved in pre-mRNA splicing. This protein is
CC associated with sn-RNP U2. It binds stem loop IV of U2 snRNA only
CC in presence of the A protein.
CC -1- SUBUNIT: Present in a spliceosome complex assembled in vitro, and
CC composed of SNRBP2, HPR8BP and CRNL1.
CC -1- DISEASE: Patients with systemic lupus erythematosus produces
CC antibodies which interact with sn-RNP proteins.
CC -1- SIMILARITY: BELONGS TO THE U1 A/B FAMILY.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: M15841; AA36796.1; -
 DR EMBL: AL034428; CAB36777.2; -
 DR PIR: A25910; A25910.
 DR PDB: 1A9N; 23-SEP-98.
 DR Genew: HGNC:1155; SNRPB2.
 DR GK: P08579; -

DR MIM: 603520; -
 DR GO: GO:0005686; C:snRNP U2; TAS.
 DR GO: GO:0008248; F:pre-mRNA splicing factor activity; TAS.
 DR GO: GO:0006371; P:mRNA splicing; TAS.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; 1.

KW mRNA processing; mRNA splicing; Nuclear protein; RNA-binding;
 KW Ribonucleoprotein; Repeat; Systemic lupus erythematosus; 3D-structure.
 FT DOMAIN 7 86 RNA-BINDING (RRM) 1.
 FT DOMAIN 151 225 RNA-BINDING (RRM) 2.
 FT STRAND 8 12
 FT TURN 16 17
 FT HELIX 20 32
 FT TURN 33 34
 FT STRAND 37 41
 FT TURN 46 50
 FT STRAND 52 56
 FT HELIX 59 68
 FT TURN 70 71
 FT STRAND 73 74
 FT TURN 75 76
 FT STRAND 77 78
 FT STRAND 80 83
 FT HELIX 89 95
 SQ SEQUENCE 225 AA; 25486 MW; 67C949C7E1A92A CRC64;

Query Match 100.0%; Score 26; DB 1; Length 225;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
 DB 72 FPFY 75

RESULT 15
 PRR2_TOBAC STANDARD; PRT; 226 AA.
 ID PRR2_TOBAC
 AC P07052;
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pathogenesis-related protein R minor form precursor (PR-R) (PROB12)
 DE (Thaumatin-like protein E2).
 OS Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
 CC NCB1 TaxID=4097;
 CC NCB1
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86230894; Pubmed=3713832;
 RA Cornelissen B.J.C.; Hooft van Hujsuijnen R.A.M.; Bol J.F.;
 RT "A tobacco mosaic virus-induced tobacco protein is homologous to the
 RL sweet-tasting protein thaumatin.";
 RN Nature 321:531-532(1986).
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Samsun NN;
 RA van Kan J.A.L.; van de Rhee M.D.; Zuidema D.; Cornelissen B.J.C.;
 RA Bol J.F.;
 RT "Structure of tobacco genes encoding thaumatin-like proteins.";
 RL Plant Mol. Biol. 12:153-155(1989).
 CC
 CC -1- SUBCELLULAR LOCATION: Vacuolar.
 CC
 CC -1- MISCELLANEOUS: PR PROTEINS ARE ACID-SOLUBLE, PROTEASE-RESISTANT
 CC PROTEINS WHICH ACCUMULATE IN THE INTERCELLULAR SPACES OF MANY
 CC PLANTS AS A RESULT OF THE HYPERSENSITIVE REACTION TO A PATHOGEN.
 CC
 CC -1- MISCELLANEOUS: PR-R EXISTS AS TWO ISOFORMS IN TOBACCO, A MAJOR AND
 CC A MINOR FORM.
 CC
 CC -1- SIMILARITY: Belongs to the thaumatin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X03913; CA27548.1; -
 DR EMBL: X15223; CA33292.1; -
 DR PIR: JH0231; JH0231.
 DR HSSP: P25871; 1AUN.
 DR InterPro: IPR001938; Thaumatin.
 DR Pfam: PF00314; thaumatin; 1.
 DR PRINTS: PR00347; THAUMATIN.
 DR ProDom: PD001321; Thaumatin; 1.
 DR SMART: SM00205; THN; 1.
 DR PROSITE: PS00316; THAUMATIN; 1.
 DR Plant defense; Pathogenesis-related protein; Signal.

FT SIGNAL 1 25
 FT CHAIN 26 226
 FT
 FT DISULFID 34 225
 FT DISULFID 75 85
 FT DISULFID 90 96
 FT DISULFID 140 214
 FT DISULFID 145 197
 FT DISULFID 153 163
 FT DISULFID 167 176
 FT DISULFID 177 184
 SQ SEQUENCE 226 AA; 24551 MW; DC9E7D9EC1813AC2 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 226;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPFY 4
 DB 7 FPFY 10

Search completed: August 20, 2003, 12:34:58
 Job time : 3.62918 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 7.56627 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-31
Perfect score: 26
Sequence: 1 PPFY 4

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	32	16	08DY47 streptococ
2	26	100.0	43	16	08F552 leptospira
3	26	100.0	45	16	09KQ98 vibrio chol
4	26	100.0	46	16	08G199 brucella su
5	26	100.0	50	17	097MA5 sulfolobus
6	26	100.0	58	2	048577 leptospira
7	26	100.0	60	13	09DGG1 gallus galli
8	26	100.0	68	16	050765 borrelia bu
9	26	100.0	69	10	08LB21 arabidopsis
10	26	100.0	70	4	08WXT4 homo sapien
11	26	100.0	70	5	09VFX9 dirosophila
12	26	100.0	72	1	054321 sulfolobus
13	26	100.0	73	5	021411 caenorhabdi
14	26	100.0	76	10	09LQ57 arabidopsis
15	26	100.0	76	12	08JX72 sen virus.
16	26	100.0	77	16	08YV55 anabaena sp

17	26	100.0	85	2	048574	048574 leptospira
18	26	100.0	88	6	08MIM2	08MIM2 bos taurus
19	26	100.0	92	17	097UJ2	097UJ2 sulfolobus
20	26	100.0	97	7	078128	078128 homo sapien
21	26	100.0	97	17	097VR9	097VR9 sulfolobus
22	26	100.0	99	16	097LJ8	097LJ8 clostridium
23	26	100.0	101	11	09CW35	09CW35 mus musculu
24	26	100.0	103	17	09HH26	09HH26 halobacteri
25	26	100.0	105	17	097XJ1	097XJ1 sulfolobus
26	26	100.0	106	7	078127	078127 homo sapien
27	26	100.0	113	5	09N923	09N923 typanosoma
28	26	100.0	113	11	08BEP9	08BEP9 mus musculu
29	26	100.0	117	8	009333	009333 nesticus si
30	26	100.0	118	5	08IS58	08IS58 ctenocephal
31	26	100.0	119	3	013530	013530 saccharomyc
32	26	100.0	123	6	095L85	095L85 oryctolagus
33	26	100.0	123	16	031793	031793 bacillus su
34	26	100.0	124	10	040395	040395 nicotiana g
35	26	100.0	124	16	08ED90	08ED90 shewanella
36	26	100.0	125	17	08TR09	08TR09 methanobarc
37	26	100.0	128	17	097US8	097US8 sulfolobus
38	26	100.0	132	11	08C3N7	08C3N7 mus musculu
39	26	100.0	135	17	08UG2	08UG2 arabidopsis
40	26	100.0	135	17	08U294	08U294 pyrococcus
41	26	100.0	136	11	08C329	08C329 mus musculu
42	26	100.0	136	16	0988W8	0988W8 rhizobium l
43	26	100.0	141	5	024953	024953 geodia cydo
44	26	100.0	142	3	042959	042959 schizosacch
45	26	100.0	147	10	09M7Y1	09M7Y1 arabidopsis

ALIGNMENTS

RESULT 1

08DY47 ID Q8DY47 PRELIMINARY; PRT; 32 AA.
AC 08DY47;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SAG1646.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Masiugnani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Jacobini E.T., Brettoni C., Galli G., Martini F., Vergi F., Malone D.,
RA Rinaldo D., Rappelli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL: AE014265; AAN0510.1; -
DR TIGR: SAG1646; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 32 AA: 3831 MW: 2020BF60F8BD8AD CRC64;

Query Match 100.0%; Score 26; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPFY 4

Db 5 FFY 8

RESULT 2

08F552 PRELIMINARY; PRT; 43 AA.
 AC 08F552;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN LA1834.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar 1a1;
 RA Ren 5.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011358; AAN49033.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 43 AA; 5278 MM; FAC6704968DB0E07 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
 Db 28 FFY 31

RESULT 3

09K098 PRELIMINARY; PRT; 45 AA.
 ID 09K098;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical protein VC2102.
 GN VC2102.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Esmolaera M.D., Vamathevan J., Bass S., Qin H., Diragol I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RT Nature 406:477-483(2000)
 RL EMBL; AE004284; AAF95248.1; -
 DR TIGR; VC2102; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 45 AA; 5437 MM; FE258F385A7CA5FD CRC64;

Query Match 100.0%; Score 26; DB 16; Length 45;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
 Db 26 FFY 29

RESULT 4

08G199 PRELIMINARY; PRT; 46 AA.
 ID 08G199;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN BR0820.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=2247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014386; AAN29749.1; -
 DR TIGR; BR0820; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 46 AA; 5493 MM; FE948C926ADDF65 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FFY 4
 Db 26 FFY 29

RESULT 5

097WAS PRELIMINARY; PRT; 50 AA.
 ID 097WAS;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE First ORF in transposon ISC1058.
 GN SSO10228.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus p2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; AE006835; AAK42483.1; -
 KW Complete proteome.
 SQ SEQUENCE 50 AA; 6354 MM; OEDF90C4DEFD27185 CRC64;

Query Match 100.0%; Score 26; DB 17; Length 50;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
Db 25 FPFY 28

RESULT 6

048577 PRELIMINARY; PRT; 58 AA.
AC 048577;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORFC.
GN ORFC.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96118697; PubMed=7496528;
RA Boursaux-Bude C., Saint Girons I., Zuercher R.;
RT "I51500, an IS3-like element from Leptospira interrogans.";
RL Microbiology 141:2165-2173(1995).
DR EMBL, U13013; AAA88920.1;
SQ SEQUENCE 58 AA; 7349 MW; D35019E9FEA8500E CRC64;

Query Match 100.0%; Score 26; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
Db 51 FPFY 54

RESULT 7

09D6G1 PRELIMINARY; PRT; 60 AA.
ID 09D6G1;
AC 09D6G1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Winged helix protein CWH-6 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Freyaldenhoven B.S., Freyaldenhoven M.P., Jaccovoni J.S., Vogt P.K.;
RT "Oncogenic potential of chicken winged helix proteins.";
RL Submitted (SEP-1995) to the EMBL/Genbank/DBD databases.
DR EMBL: U37277; AAF97843.1;
DR HSSP: Q63245; ZHFN.
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR ProDom: PD000425; TF_Fork_head; 1.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 60 AA; 7351 MW; B20EDB371F3AD507 CRC64;

Query Match 100.0%; Score 26; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
Db 9 FPFY 12

RESULT 8

050765 PRELIMINARY; PRT; 68 AA.
ID 050765;
AC 050765;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein BbJ10.
GN BbJ10.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,
RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D., Hanson M.,
RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Uterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horel K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL: AE000787; AAC66097.1;
DR TIGR: BbJ10;
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 68 AA; 7927 MW; 03E39BDB6E24C09A CRC64;

Query Match 100.0%; Score 26; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
1111
Db 45 FPFY 48

RESULT 9

08LBZ1 PRELIMINARY; PRT; 69 AA.
ID 08LBZ1;
AC 08LBZ1;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troupkan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Broeyer V., Troupkan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/Genbank/DBD databases.

```

DR EMBL: AY086916; AAM64480.1; -.
KW Hypothetical protein.
SQ SEQUENCE 69 AA; 7446 MW; 9D7D16D302F6B973 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
   ||||
Db 23 FFY 26

RESULT 10
O8WXT4 PRELIMINARY; PRT; 70 AA.
ID O8WXT4
AC O8WXT4;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Winged helix transcription factor delta (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Freyaldenhoven B.S., Fried C., Wielckens K.;
RA "Homo sapiens myeloid factor-delta mRNA.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF343007; AAL73344.1; -.
DR InterPro: IPR001766; TF_Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR PRODOM: PD000425; TF_Fork_head; 1.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
FT NON_TER 1
FT TER 70
SQ SEQUENCE 70 AA; 8355 MW; 96487388A0DD80B2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
   ||||
Db 21 FFY 24

RESULT 11
Q9VFX9 PRELIMINARY; PRT; 70 AA.
ID Q9VFX9
AC Q9VFX9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Cg11686 protein (RH19248P).
GN Cg11686 OR BC DNA: RH19248.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RX MEDLINE:20196006; Pubmed-10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokoyva D., Botchan M.R., Bouck J., Brinkstein P., Brottler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RC Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Flise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003699; AAF54916.1; -.
DR EMBL: AT070684; AAL48155.1; -.
DR Flybase: FBgn0040551; CG11686.
SQ SEQUENCE 70 AA; 8200 MW; 57DDP90DE968256 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 5; Length 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFY 4
   ||||
Db 55 FFY 58

RESULT 12
Q54321 PRELIMINARY; PRT; 72 AA.
ID Q54321
AC Q54321;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-JUN-1998 (TREMblrel. 06, Last annotation update)
OS Sulfolobus islandicus.
OG Sulfolobus islandicus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=43080;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-RENH1;
RX MEDLINE-96271189; PubMed-8700967;
RA Keeling P.J., Klenk H.P., Singh R.K., Feeley O., Schlieper C.,
RA Zillig W., Doolittle W.F., Sensen C.W.;
RT "Complete nucleotide sequence of the Sulfolobus islandicus multicopy
RT plasmid pRM1."
RL plasmid 35:141-144(1996).
DR EMBL; U36383; AAC44108.1; -.
KW plasmid.
SQ SEQUENCE 72 AA; 8412 MW; 7C48074271256FB5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 1; Length 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FFY 4
Db 5 FFY 8

RESULT 13
ID 021411 PRELIMINARY; PRT; 73 AA.
AC 021411;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical 8.3 kDa protein.
GN K10B3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Galling S.;
RT "The sequence of C. elegans cosmid K10B3."
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submision."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49941; AAB53870.1; -.
DR WormPep; K10B3.1; CE07364.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8309 MW; 49BE84CA4AE53901 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 5; Length 73;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FFY 4
Db 52 FFY 55

RESULT 14
ID 091Q57 PRELIMINARY; PRT; 76 AA.
AC 091Q57;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

```

```

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE T30E16.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN 11
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Dunn P., Khan S., Kim C., Shin P.,
RA Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A., Hansen N.F.,
RA Hultzer L., Kremetskaia I., Lenz C., Li J., Liu S., Luros S.,
RA Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G., Davis R.W.,
RA Federspiel N.A., Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T30E16 from chromosome
RT 1."
RN Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN 14
RP SEQUENCE FROM N.A.
RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chou J., Choi E., Lam B.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharbeky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick N.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009317; AAF79751.1; -.
DR InterPro; IPR000834; zn_cardocept.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
SQ SEQUENCE 76 AA; 8899 MW; FDCDD69597E90B8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FFY 4
Db 54 FFY 57

RESULT 15
ID 08UX72 PRELIMINARY; PRT; 76 AA.
AC 08UX72;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE ORF1 protein (Fragment).
DE SEN virus.
OS Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_Taxid=136966;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-SEN-D;
RA Tanaka Y., Pfeiffer R., Yeo A.E., Mizokami M., Edlin B.R.,
RA O'Brien T.R., Alter H.J., Shih J.W.;
RT "PCR-probe capture hybridization assay and statistical mixture
RT modeling for SEN virus prevalence estimation."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB075270; BAB97340.1; -.
DR InterPro; IPR004219; TTVirus_unk.
DR Pfam; PF02956; TT_ORF1; 1.
FT NON_TER 1 76
FT

```

SQ SEQUENCE 76 AA; 8656 MW; C62AF6E245301A13 CRC64;

Query Match 100.0%; Score 26; DB 12; Length 76;

Best Local Similarity 100.0%; Pred. No. 2.6e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FPFY 4
||||

DB 49 FPFY 52

Search completed: August 20, 2003, 12:40:30
Job time : 10.6774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 : Search time 13.1928 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-32
Perfect score: 48
Sequence: 1 NGWYPM 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	87.5	562	22	ABG03291
2	40	83.3	403	21	AAB26455
3	40	83.3	414	21	AAB20924
4	39	81.2	853	22	AAG55229
5	38	79.2	39	22	ABG59477
6	38	79.2	39	22	ABBA4094
7	38	79.2	39	22	ABR26986
8	38	79.2	39	22	AAK55117
9	38	79.2	39	22	AAW77824

10	38	79.2	39	22	AAM21726
11	38	79.2	39	22	AAM38042
12	38	79.2	39	22	ABG46855
13	38	79.2	81	21	AAG48003
14	38	79.2	83	21	AAG07472
15	38	79.2	99	22	AAU47631
16	38	79.2	117	24	ABP78660
17	38	79.2	146	21	AAV99453
18	38	79.2	146	22	AAU29211
19	38	79.2	146	22	AAG74396
20	38	79.2	146	22	AAB66202
21	38	79.2	146	24	ABU71299
22	38	79.2	146	24	ABU65756
23	38	79.2	146	24	ABU66089
24	38	79.2	146	24	ABU67593
25	38	79.2	146	24	ABU65451
26	38	79.2	146	24	ABU55857
27	38	79.2	146	24	ABU56123
28	38	79.2	146	24	ABU57118
29	38	79.2	146	24	ABU10697
30	38	79.2	337	21	AAB07589
31	38	79.2	343	21	AAG47999
32	38	79.2	452	21	AAG53566
33	38	79.2	454	21	AAG53565
34	38	79.2	557	21	AAG53564
35	38	79.2	557	23	ABR90965
36	37.5	78.1	467	21	AAG41264
37	37.5	78.1	544	21	AAG41263
38	37.5	78.1	573	21	AAG41262
39	37	77.1	12	23	ABR06171
40	37	77.1	38	22	ABG53484
41	37	77.1	38	22	ABR38629
42	37	77.1	38	22	ABR23722
43	37	77.1	38	22	AAW59259
44	37	77.1	38	22	AAW71800
45	37	77.1	38	22	AAM19313

ALIGNMENTS

RESULT 1
ABG03291
ID ABG03291 standard; Protein: 562 AA.
XX
XX
AC ABG03291;
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE Novel human diagnostic protein #3282.
XX
XX
DE Human: chromosome mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200175067-A2.
XX
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
XX
XX
PR 23-AUG-2000; 2000US-0649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Dimaac RT, Liu C, Tang YT;
XX
XX
DR WPI: 2001-639362/73.
XX
XX
DR N-PSDB; AAG7478.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity
PS	
XX	Claim 20; SEQ ID No 33650; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG3037 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 562 AA;
	Query Match 87.5%; Score 42; DB 22; Length 562;
	Best Local Similarity 100.0%; Pctd. No. 85;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	2 GMYPW 6
Db	367 GMYPW 371
RESULT 2	
ID	AAB26455 standard; Protein; 403 AA.
XX	AAB26455
AC	AAB26455;
XX	
DT	23-FEB-2001 (first entry)
XX	
DE	Drosophila melanogaster odorant receptor DOR83.
XX	
KW	Odorant receptor; fruit fly; DOR83; odour recognition; pest control.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO2000050566-A2.
XX	
PD	31-AUG-2000.
XX	
PF	25-FEB-2000; 2000WO-US04995.
XX	
PR	25-FEB-1999; 99US-0257706.
XX	
PA	(UYCO) UNITV COLUMBIA NEW YORK.
XX	
PI	Vosshall LB, Amrein HO, Axel R;
XX	
PI	WPI; 2000-572081/53.
DR	N-PSDB; AAA94853.
XX	
PT	Novel nucleic acid encoding an insect odorant receptor, for identifying
PT	modulator compounds that are useful in controlling pest population
XX	
XX	disclosure; Page 86; 176pp; English.

CC	The present sequence is the previously identified Drosophila melanogaster
CC	odorant receptor DOR83. The odorant genes and proteins, such as those
CC	provided by the invention, are useful as they aid in the study of the
CC	olfactory organ in mammals, as well as aiding the understanding of the
CC	link between odour recognition and behaviour in insects. They also enable
CC	the identification of compounds capable of activating and inhibiting the
CC	receptors, allow the control of pest populations via the use of alarm
CC	odour ligands and via the use of ligands which interfere with the
CC	interaction between odorant ligands and receptors associated with the
CC	fertility.
XX	
SQ	Sequence 403 AA;
OY	
	1 NGWYPW 6 I I I I I
Db	176 NTWYPW 181
RESULT 3	
AAB20924	
ID	AAB20924 standard; Protein; 414 AA.
XX	
AC	AAB20924;
XX	
DT	06-DEC-2000 (first entry)
XX	
DE	Drosophila odorant receptor DOR 69F.2.
KX	
KM	Odorant receptor; Drosophila; Olfactory receptor;
KW	G protein-coupled receptor; GPCR superfamily; transgenic insect;
KM	Insect behaviour modification; pest control; pollinator attraction;
KW	biosensor; odour detection; odour identification; apiculture.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200043410-A2.
XX	
PD	27-JUL-2000.
PX	
PF	25-JAN-2000; 2000MO-USO1823.
XX	
PR	25-JAN-1999; 99US-0117132.
XX	
PA	(UYYA) UNIV YALE.
XX	
PI	Carlson JR, Kim J, Clyne PJ, Mair CG;
XX	
DR	WPL: 2000-543246/49.
DR	N-PSDB: AAA72259.
XX	
	New nucleic acid encoding a Drosophila olfactory receptor, useful for
PT	identifying modulating agents -
XX	
PS	Claim 12; Page 184-185; 303pp; English.
XX	
CC	Sequences AAB20901-B20949 represent Drosophila melanogaster odorant
CC	receptors. These proteins function as olfactory receptors, and
CC	are thought to be members of the G protein-coupled receptor
CC	(GPCR) superfamily, which is characterised by the presence of 7
CC	transmembrane helices. Nucleic acids encoding the Drosophila odorant
CC	receptors may be used to generate expression constructs, host cells
CC	containing such constructs, and transgenic insects. Cells which express
CC	the odorant receptor genes may be used in methods to identify agents
CC	which modulate expression of these genes, and in methods to identify
CC	receptor binding partners. The Drosophila odorant receptor nucleic acids
CC	may also be used to identify corresponding genes in other insects,
CC	such as those which damage crops or transmit disease. The odorant
CC	receptor proteins may be used to identify agents which modulate their
CC	activity, to identify binding partners, as antigens to raise antibodies,
CC	

CC and in methods to modify insect behaviour. The proteins may be also
 CC be used in methods of behaviour modification. Such methods may be used
 CC to study or modify insect behaviour in response to odorants such as
 CC pheromones. Modification of insect behaviour has a wide range of
 CC applications, such as in pest control (e.g., by disrupting the feeding
 CC or mating behaviours of pest species), or for enhancing plant
 CC pollination (by attracting pollinator species). Odorant receptor proteins
 CC and/or nucleotides may also be used to identify appetite suppressants, to
 CC trap odours of a specific type, as biosensors for the detection of
 CC explosives, drugs, perfumes or pollutants, and in apiculture to modify
 CC the behaviour of bees, for example, to increase the production of royal
 CC jelly.

XX Sequence 414 AA:

Query Match 83.3%; Score 40; DB 21; Length 414;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 | | | | |
 Db 176 NTWYPM 181

RESULT 4
 AAG65229
 ID AAG65229 standard; Protein: 853 AA.

XX AAG65229;

XX 20-NOV-2001 (first entry)

DE Mandrill immunodeficiency virus SIM27 ENV protein.

KM Simian immunodeficiency virus; HIV; gag; env; pol; mandrill;
 antibody detection.

OS Simian immunodeficiency virus.

PN NO200003889-A.

PD 05-FEB-2001.

PF 28-JUL-2000; 2000NO-0003889.

XX 03-AUG-1999; 99DE-1036003.

PA (DADE-) DADE BEHRING MARBURG GMBH.

PI Guertler LG, Kaptue L, Zekeng L, Deloko YBD, Hauser HP;

DR WPI: 2001-201061/20.

PT New mandrill immunodeficiency virus SIM27 useful for detecting
 antibodies directed against the virus -

PS Example 3: Page 27; 47pp; Norwegian.

CC The present invention relates to a new mandrill immunodeficiency
 CC virus SIM27, which can be used to detect antibodies directed against
 CC the virus. The present sequence is the ENV protein described in the
 CC exemplification of the invention.

XX Sequence 853 AA:

Query Match 81.2%; Score 39; DB 22; Length 853;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 | | | | |
 Db 416 NWYPM 421

RESULT 5
 ABG59477
 ID ABG59477 standard; Peptide: 39 AA.

XX ABG59477;

XX 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 38125.

KM Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 hypercholesterolaemia; coronary heart disease.

KW Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00664.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-48898/53.

XX Human genome-derived single exon nucleic acid probes useful for
 analysing gene expression in human adult liver -

PT Claim 27; SEQ ID No 38125; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which

CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.

CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 39 AA:

Query Match 79.2%; Score 38; DB 22; Length 39;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 : | : | : |
 Db 10 SGWHPM 15

RESULT 6
 ABB44094
 ID ABB44094 standard; Peptide: 39 AA.

XX ABB44094;

DT 04-FEB-2002 (first entry)
 XX Peptide #11600 encoded by human foetal liver single exon probe.
 DE
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KM
 XX Homo sapiens.
 OS
 XX WO200157277-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 PR
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver.
 PT
 XX
 PS Claim 27; SEQ ID NO 36729; 639pp + sequence listing; English.
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 39 AA;
 Query Match 79.2%; Score 38; DB 22; Length 39;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps -0;
 OY 1 NGMYPW 6
 :||:|
 Db 10 SGWHPW 15
 RESULT 7
 ABB26986
 ID ABB26986 standard; Protein; 39 AA.
 AC ABB26986;
 XX
 XX 23-JAN-2002 (first entry)
 DT
 XX Protein #8985 encoded by probe for measuring heart cell gene expression.
 DE
 XX Human; gene expression; heart; microarray; vascular system;
 KM cardiovascular disease; hypertension; cardiac arrhythmia;
 KM congenital heart disease.
 KW
 XX Homo sapiens.
 OS
 XX WO200157274-A2.
 PN
 XX 09-AUG-2001.
 PD

XX
 PF 30-JAN-2001; 2001WO-US00666.
 PR
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488699/53.
 DR
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PT
 XX
 PS Claim 15; SEQ ID NO 28756; 530pp; English.
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 39 AA;
 Query Match 79.2%; Score 38; DB 22; Length 39;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NGMYPW 6
 :||:|
 Db 10 SGWHPW 15
 RESULT 8
 AAM65117
 ID AAM65117 standard; Protein; 39 AA.
 AC AAM65117;
 XX
 XX 05-NOV-2001 (first entry)
 DT
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 37222.
 DE
 XX Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 KW
 XX Homo sapiens.
 OS
 XX WO200157275-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00667.
 PR
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 37222; 650bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 39 AA;
 XX
 OY Query Match 79.2%; Score 38; DB 22; Length 39;
 DB Best Local Similarity 66.7%; Pred. No. 25;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NGWYPPW 6
 :||:|
 10 SGWHPW 15
 DB
 RESULT 9
 AAM77824
 ID AAM77824 standard; Protein; 39 AA.
 XX
 AC AAM77824;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38130.
 XX
 DE Human bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001MO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 38130; 658bp + Sequence Listing; English.
 XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 39 AA;
 XX
 OY Query Match 79.2%; Score 38; DB 22; Length 39;
 DB Best Local Similarity 66.7%; Pred. No. 25;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NGWYPPW 6
 :||:|
 10 SGWHPW 15
 DB
 RESULT 10
 AAM21726
 ID AAM21726 standard; Protein; 39 AA.
 XX
 AC AAM21726;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #8160 encoded by probe for measuring cervical gene expression.
 DE
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001MO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID NO 26552; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 39 AA;
 XX
 OY Query Match 79.2%; Score 38; DB 22; Length 39;
 XX

Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
:|:|:|
Db 10 SGWHPW 15

RESULT 11

AAM38042
ID AAM38042 standard; Protein; 39 AA.

AC AAM38042;

DT 17-OCT-2001 (first entry)

DE Peptide #12079 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;

KM genetic disorder.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for

PS analyzing gene expression in human placenta.

PS Claim 27; SEQ ID No 38311; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP:

CC see AAI31315-AA157546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for

CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.

SQ Sequence 39 AA;

OY Query Match 79.2%; Score 38; DB 22; Length 39;

Db Best Local Similarity 66.7%; Pred. No. 25;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 12

ABG46855

ID ABC46855 standard; Peptide; 39 AA.

AC ABC46855;

DT 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 36520.

DE Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

OS WO200186003-A2.

PN 15-NOV-2001.

PD 30-JAN-2001; 2001WO-US000665.

PF 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to

PS measure gene expression in human lung samples -

PS Claim 27; SEQ ID No 36520; 654bp; English.

CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridize at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
XX
SQ Sequence 39 AA;

Query Match 79.28; Score 38; DB 23; Length 39;
Best Local Similarity 66.78; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
:1:1:1
DB 10 SGWHPW 15

RESULT 13
AAG48003
ID AAG48003 standard; Protein: 81 AA.

XX
AC AAG48003;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 60571.

DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX
EP1033405-A2.

PD 06-SEP-2000.

XX
PE 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.

XX
PR 05-MAR-1999; 99US-0123160.

XX
PR 09-MAR-1999; 99US-0123548.

XX
PR 23-MAR-1999; 99US-0125768.

XX
PR 25-MAR-1999; 99US-0126264.

XX
PR 29-MAR-1999; 99US-0126785.

XX
PR 01-APR-1999; 99US-0127462.

XX
PR 06-APR-1999; 99US-0128234.

XX
PR 16-APR-1999; 99US-0128714.

XX
PR 19-APR-1999; 99US-0129845.

XX
PR 21-APR-1999; 99US-0130077.

XX
PR 23-APR-1999; 99US-0130510.

XX
PR 23-APR-1999; 99US-0130891.

XX
PR 28-APR-1999; 99US-0131449.

XX
PR 30-APR-1999; 99US-0132048.

XX
PR 04-MAY-1999; 99US-0132484.

XX
PR 05-MAY-1999; 99US-0132486.

XX
PR 06-MAY-1999; 99US-0132486.

XX
PR 07-MAY-1999; 99US-0132863.

XX
PR 11-MAY-1999; 99US-0134256.

XX
PR 14-MAY-1999; 99US-0134218.

XX
PR 14-MAY-1999; 99US-0134219.

XX
PR 14-MAY-1999; 99US-0134221.

XX
PR 14-MAY-1999; 99US-0134370.

XX
PR 18-MAY-1999; 99US-0134768.

XX
PR 19-MAY-1999; 99US-0134941.

XX
PR 20-MAY-1999; 99US-0135124.

XX
PR 21-MAY-1999; 99US-0135353.

XX
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 20-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145090.

PR 22-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145214.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 27-JUL-1999; 99US-0145921.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159320.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match
Best Local Similarity 79.2%; Score 38; DB 21; Length 81;
Matches 4; Conservative 80.0%; Pred. No. 52;
1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GWYPM 6
Db 10 GWYPM 14
RESULT 14
AAG07472
ID AAG07472 standard; Protein; 83 AA.
XX
AC AAG07472;
XX
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4640.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149476.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155138.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157735.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160810.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 38; DB 21; Length 83;
Best Local Similarity 80.0%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWYPM 6
11:11
Db 12 GWYPM 16

QY 2 GWYPM 6
11:11
Db 56 GWYPM 60

Search completed: August 20, 2003, 12:33:46
Job time : 14.1928 secs

RESULT 15
AAU47631
ID AAU47631 standard; Protein; 99 AA.
XX
AC AAU47631;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #8527.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN W0200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59539.
XX
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 8826; 1069pp; English.
XX
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 99 AA;

Query Match 79.2%; Score 38; DB 22; Length 99;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 4.15663 Seconds
(without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-32

Perfect score: 48
Sequence: 1 NCWYPM 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	87.5	115	US-09-252-991A-16759	Sequence 16759, A
2	39	81.2	853	US-09-625-972-25	Sequence 25, Appl
3	38	79.2	236	US-09-252-991A-28172	Sequence 28172, A
4	37.5	78.1	555	US-09-252-991A-20604	Sequence 20604, A
5	36	75.0	13	US-08-915-314-34	Sequence 34, Appl
6	36	75.0	13	US-09-030-619-35	Sequence 35, Appl
7	36	75.0	13	US-09-667-486-34	Sequence 34, Appl
8	36	75.0	190	US-09-355-166-7	Sequence 7, Appl
9	36	75.0	334	US-08-279-058B-10	Sequence 10, Appl
10	36	75.0	334	US-08-828-323-10	Sequence 10, Appl
11	36	75.0	679	US-09-198-452A-1132	Sequence 1132, Ap
12	36	75.0	684	US-09-134-078-62	Sequence 62, Appl
13	36	75.0	703	US-09-252-991A-27442	Sequence 27442, A
14	35	72.9	212	US-08-860-255A-4	Sequence 4, Appl
15	35	72.9	273	US-08-484-905-118	Sequence 118, App
16	35	72.9	273	US-08-481-985B-118	Sequence 118, App
17	35	72.9	273	US-08-370-476-118	Sequence 118, App
18	35	72.9	273	US-09-311-784A-30	Sequence 30, Appl
19	35	72.9	288	US-08-160-544A-6	Sequence 6, Appl
20	35	72.9	345	US-09-120-365-73	Sequence 73, Appl
21	35	72.9	345	US-09-515-039-73	Sequence 73, Appl
22	35	72.9	345	US-08-827-171B-7	Sequence 7, Appl
23	34	70.8	6	US-08-871-561-23	Sequence 23, Appl
24	34	70.8	6	US-09-321-932B-23	Sequence 23, Appl
25	34	70.8	76	US-09-663-600A-115	Sequence 115, App
26	34	70.8	76	US-09-663-600A-209	Sequence 209, App
27	34	70.8	82	US-09-663-600A-117	Sequence 117, App

28	34	70.8	92	US-09-663-600A-211	Sequence 211, App
29	34	70.8	184	US-09-252-991A-26087	Sequence 26087, A
30	34	70.8	197	US-09-328-352-5731	Sequence 5731, Ap
31	34	70.8	218	US-08-985-526-1	Sequence 1, Appl
32	34	70.8	239	PCT-US93-01652-1	Sequence 1, Appl
33	34	70.8	270	US-09-252-991A-19648	Sequence 19648, A
34	34	70.8	312	US-09-267-117-21	Sequence 21, Appl
35	34	70.8	314	US-09-583-545-15	Sequence 15, Appl
36	34	70.8	335	US-08-736-915-2	Sequence 2, Appl
37	34	70.8	360	US-09-252-991A-26691	Sequence 26691, A
38	34	70.8	433	US-09-252-991A-25248	Sequence 25248, A
39	34	70.8	441	US-08-985-526-3	Sequence 3, Appl
40	34	70.8	575	US-09-252-991A-20649	Sequence 20649, A
41	34	70.8	904	US-09-252-991A-19257	Sequence 19257, A
42	34	70.8	980	US-08-220-151-5	Sequence 5, Appl
43	34	70.8	980	US-08-413-118-5	Sequence 5, Appl
44	34	70.8	980	US-08-473-446-5	Sequence 5, Appl
45	34	70.8	1170	US-08-313-288B-20	Sequence 20, Appl

ALIGNMENTS

```

RESULT 1
US-09-252-991A-16759
: Sequence 16759, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: MARC J. RUBENFIELD et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 16759
: LENGTH: 115
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-16759

Query Match      87.5% Score 42; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GWYPM 6
Db      56 GWYPM 60

RESULT 2
US-09-625-972-25
: Sequence 25, Application US/09625972
: Patent No. 656513
: GENERAL INFORMATION:
: APPLICANT: GUERTLER, Lutz Gerhard
: APPLICANT: HAUSER, Hans Peter
: APPLICANT: DONGMO DELOKO, Yvette Beatrice
: APPLICANT: ZERKENG, Leopold
: APPLICANT: KAPPU, Lazare
: TITLE OF INVENTION: LENTIVIRUS FROM THE GROUP OF IMMUNODEFICIENCY VIRUSES OF DRILL
: FILE REFERENCE: 067595/0106
: CURRENT APPLICATION NUMBER: US/09/625,972
: CURRENT FILING DATE: 2000-07-29
: PRIOR APPLICATION NUMBER: DE 199 36 003.0
: PRIOR FILING DATE: 1999-08-03
: NUMBER OF SEQ ID NOS: 57
: SOFTWARE: PatentIn version 3.0

```

SEQ ID NO 25
LENGTH: 853
TYPE: PRT
ORGANISM: SIV - Viral
US-09-625-972-25

Query Match
Best Local Similarity 81.2%; Score 39; DB 4; Length 853;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 416 NWYPM 421

RESULT 3
US-09-252-991A-28172
Sequence 28172, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28172
LENGTH: 236
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28172

Query Match
Best Local Similarity 79.2%; Score 38; DB 4; Length 236;
Best Local Similarity 80.0%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWYPM 6
DB 13 GWYPM 17

RESULT 4
US-09-252-991A-20604
Sequence 20604, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20604
LENGTH: 555
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (202)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20604

Query Match
78.1%; Score 37.5; DB 4; Length 555;

Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 NGWYPM 6
DB 250 NGWYPM 256

RESULT 5
US-08-915-314-34
Sequence 34, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-34

Query Match
Best Local Similarity 75.0%; Score 36; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
DB 7 WYPM 10

RESULT 6
US-09-030-619-35
Sequence 35, Application US/09030619B
Patent No. 6503881
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolizidin Analogue
US-09-030-619-35

Query Match 75.0%; Score 36; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
Db 7 WYPM 10

RESULT 7
US-09-667-486-34
; Sequence 34, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; West, Michael H.P.
; Krieger, Timothy J.
; Taylor, Robert
; Eflie, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING ANALOGUES OF INDOLIZIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/667,486
; FILING DATE: 22-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6538106Lenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-667-486-34

Query Match 75.0%; Score 36; DB 4; Length 13;

Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
Db 7 WYPM 10

RESULT 8
US-09-355-166-7
; Sequence 7, Application US/09355166
; Patent No. 6316241
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydroxylase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/09/355,166
; CURRENT FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Bacillus
US-09-355-166-7

Query Match 75.0%; Score 36; DB 4; Length 190;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
1111
Db 17 NHWYPM 22

RESULT 9
US-08-279-058B-10
; Sequence 10, Application US/08279058B
; Patent No. 5668004
; GENERAL INFORMATION:
; APPLICANT: Michael E. O'Donnell
; TITLE OF INVENTION: DNA POLYMERASE III
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,058B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D-1056CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-279-058B-10

Query Match 75.0%; Score 36; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
DB 3 WYPM 6

RESULT 10

US-08-828-323-10
; Sequence 10, Application US/08828323A
; Patent No. 6413753
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael
; TITLE OF INVENTION: DNA POLYMERASE III HOLOENZYME
; FILE REFERENCE: 19603/10214
; CURRENT APPLICATION NUMBER: US/08/828,323A
; CURRENT FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 10
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-828-323-10

Query Match 75.0%; Score 36; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
DB 3 WYPM 6

RESULT 11

US-09-198-452A-1132
; Sequence 1132, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO: 1132
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1132

Query Match 75.0%; Score 36; DB 4; Length 679;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
DB 28 WYPM 31

RESULT 12

US-09-134-078-62
; Sequence 62, Application US/09134078
; Patent No. 6368844
; GENERAL INFORMATION:
; APPLICANT: Bylina, Edward J.
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Gray Cary Ware & Freidenrich LLP

STREET: 4365 Executive Drive, Suite 1600

CITY: San Diego

STATE: CA

COUNTRY: USA

ZIP: 92121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134,078

FILING DATE: 13-AUG-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/949,026

FILING DATE: 10-OCT-1997

APPLICATION NUMBER: 60/056,916

FILING DATE: 06-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09010/024002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 858/677-1456

TELEFAX: 858/677-1465

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 684 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: Internal

Query Match 75.0%; Score 36; DB 4; Length 684;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
1111
DB 252 WYPM 255

RESULT 13

US-09-252-991A-27442
; Sequence 27442, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 27442
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27442

Query Match 75.0%; Score 36; DB 4; Length 705;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 1111
 Db 534 NTWPM 539

RESULT 14
 US-08-860-255A-4
 ; Sequence 4, Application US/08860255A
 ; Patent No. 6274336
 ; GENERAL INFORMATION:
 ; APPLICANT: Abdel-Meguid, Sherin
 ; APPLICANT: Desjarlais, Renee
 ; APPLICANT: Janson, Cheryl
 ; APPLICANT: Smith, Ward
 ; APPLICANT: Zhao, Baoguang
 ; TITLE OF INVENTION: Method of Inhibiting Cathepsin K
 ; FILE REFERENCE: P50574-XI
 ; CURRENT APPLICATION NUMBER: US/08/860,255A
 ; CURRENT FILING DATE: 1997-06-26
 ; PRIOR APPLICATION NUMBER: 60/008,108
 ; PRIOR FILING DATE: 1995-10-30
 ; PRIOR APPLICATION NUMBER: 60/007,473
 ; PRIOR FILING DATE: 1995-11-22
 ; PRIOR APPLICATION NUMBER: 60/008,992
 ; PRIOR FILING DATE: 1995-12-21
 ; PRIOR APPLICATION NUMBER: 60/013,748
 ; PRIOR FILING DATE: 1996-03-20
 ; PRIOR APPLICATION NUMBER: 60/013,764
 ; PRIOR FILING DATE: 1996-03-20
 ; PRIOR APPLICATION NUMBER: 60/013,747
 ; PRIOR FILING DATE: 1996-03-20
 ; PRIOR APPLICATION NUMBER: 60/017,455
 ; PRIOR FILING DATE: 1996-05-17
 ; PRIOR APPLICATION NUMBER: 60/017,892
 ; PRIOR FILING DATE: 1996-05-17
 ; PRIOR APPLICATION NUMBER: 60/020,478
 ; PRIOR FILING DATE: 1996-06-13
 ; PRIOR APPLICATION NUMBER: 60/022,047
 ; PRIOR FILING DATE: 1996-07-22
 ; PRIOR APPLICATION NUMBER: 60/023,494
 ; PRIOR FILING DATE: 1996-08-07
 ; PRIOR APPLICATION NUMBER: 60/023,742
 ; PRIOR FILING DATE: 1996-08-08
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-08-860-255A-4

Query Match 72.9%; Score 35; DB 3; Length 212;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 1111
 Db 64 NGWYPM 69

RESULT 15
 US-08-484-905-118
 ; Sequence 118, Application US/08484905
 ; Patent No. 5976551
 ; GENERAL INFORMATION:
 ; APPLICANT: Motiez, Estelle
 ; APPLICANT: Abastado, Jean-Pierre
 ; APPLICANT: Kourilsky, Philippe
 ; TITLE OF INVENTION: An Altered Major Histocompatibility
 ; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
 ; NUMBER OF SEQUENCES: 127
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &
 ADDRESSEE: Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,905
 FILING DATE: 07-JUNE-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/801,818
 FILING DATE: 05-DEC-1991
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/792,473
 FILING DATE: 15-NOV-1991
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Potter, Jane E. R.
 REGISTRATION NUMBER: 33,332
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4400
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 118:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 273 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-484-905-118

Query Match 72.9%; Score 35; DB 2; Length 273;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 1111
 Db 3 SGWYPM 8

Search completed: August 20, 2003, 12:44:22
 Job time: 5.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 6.50602 Seconds
(Without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-32

Perfect score: 48

Sequence: 1 NGWYFW 6

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	6	11 US-09-300-425B-32	Sequence 32, Appl
2	43	89.6	318	15 US-10-156-761-376	Sequence 7581, App
3	40	83.3	403	12 US-10-183-708-92	Sequence 92, Appl
4	38	79.2	39	9 US-09-864-761-42284	Sequence 42284, A
5	38	79.2	146	11 US-09-946-374-390	Sequence 390, App
6	38	79.2	146	12 US-10-015-387A-390	Sequence 390, App
7	38	79.2	146	12 US-10-006-130A-390	Sequence 390, App
8	38	79.2	146	12 US-10-199-672-376	Sequence 376, App
9	38	79.2	146	12 US-10-006-172A-390	Sequence 390, App
10	38	79.2	146	12 US-10-187-745-376	Sequence 376, App
11	38	79.2	146	12 US-10-194-457-376	Sequence 376, App
12	38	79.2	146	14 US-10-052-586-376	Sequence 376, App
13	38	79.2	146	15 US-10-174-590-376	Sequence 376, App
14	38	79.2	146	15 US-10-176-758-376	Sequence 376, App
15	38	79.2	146	15 US-10-175-737-376	Sequence 376, App

16	38	79.2	146	15 US-10-173-706-376	Sequence 376, App
17	38	79.2	146	15 US-10-175-738-376	Sequence 376, App
18	38	79.2	146	15 US-10-175-752-376	Sequence 376, App
19	38	79.2	146	15 US-10-176-482-376	Sequence 376, App
20	38	79.2	146	15 US-10-176-757-376	Sequence 376, App
21	38	79.2	146	15 US-10-176-913-376	Sequence 376, App
22	38	79.2	146	15 US-10-180-552-376	Sequence 376, App
23	38	79.2	146	15 US-10-180-557-376	Sequence 376, App
24	38	79.2	146	15 US-10-173-700-376	Sequence 376, App
25	38	79.2	146	15 US-10-174-572-376	Sequence 376, App
26	38	79.2	146	15 US-10-174-579-376	Sequence 376, App
27	38	79.2	146	15 US-10-174-582-376	Sequence 376, App
28	38	79.2	146	15 US-10-174-588-376	Sequence 376, App
29	38	79.2	146	15 US-10-175-739-376	Sequence 376, App
30	38	79.2	146	15 US-10-175-740-376	Sequence 376, App
31	38	79.2	146	15 US-10-175-743-376	Sequence 376, App
32	38	79.2	146	15 US-10-176-488-376	Sequence 376, App
33	38	79.2	146	15 US-10-176-492-376	Sequence 376, App
34	38	79.2	146	15 US-10-176-747-376	Sequence 376, App
35	38	79.2	146	15 US-10-176-750-376	Sequence 376, App
36	38	79.2	146	15 US-10-176-985-376	Sequence 376, App
37	38	79.2	146	15 US-10-176-987-376	Sequence 376, App
38	38	79.2	146	15 US-10-176-992-376	Sequence 376, App
39	38	79.2	146	15 US-10-176-993-376	Sequence 376, App
40	38	79.2	146	15 US-10-184-658-376	Sequence 376, App
41	38	79.2	146	15 US-10-176-991-376	Sequence 376, App
42	38	79.2	146	15 US-10-173-695-376	Sequence 376, App
43	38	79.2	146	15 US-10-173-697-376	Sequence 376, App
44	38	79.2	146	15 US-10-173-705-376	Sequence 376, App
45	38	79.2	146	15 US-10-174-576-376	Sequence 376, App

ALIGNMENTS

RESULT 1
US-09-300-425B-32
; Sequence 32, Application US/09300425B
; Publication No. US20030045681A1
; GENERAL INFORMATION:
; APPLICANT: NERI, Danilo
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
; OTHER INFORMATION: antibody clone
US-09-300-425B-32

Query Match 100.0%; Score 48; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 4; 4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYFW 6
|||||
Db 1 NGWYFW 6

RESULT 2

US-10-156-761-7581
; Sequence 7581, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7581
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7581

Query Match 89.6%; Score 43; DB 15; Length 318;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 106 SGWYPM 111

RESULT 3
US-10-183-708-92
; Sequence 92, Application US/10183708
; Publication No. US20030143679A1
; GENERAL INFORMATION:
; APPLICANT: VOSSHALL, LESLIE
; APPLICANT: AMREIN, HUBERT
; APPLICANT: AXEL, RICHARD
; TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
; FILE REFERENCE: 0575/58715-AA-PCT-US/JPM/ADM/BJA
; CURRENT APPLICATION NUMBER: US/10/183,708
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 09/932,227
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/04995
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/257,706
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Drosophila Melanogaster DOR83
US-10-183-708-92

Query Match 83.3%; Score 40; DB 12; Length 403;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 176 NTWYPM 181

RESULT 4
US-09-864-42284
; Sequence 42284, Application US/09864761

Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42284
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018720.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 45
; OTHER INFORMATION: EST HUMAN HIT: BB880162.1, EVALUATE 1.00e-20
; OTHER INFORMATION: SWISSPROT HIT: Q07092, EVALUATE 4.90e+00
US-09-864-42284

Query Match 79.2%; Score 38; DB 9; Length 39;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6

Db 10 SCWHPW 15

RESULT 5
US-09-946-374-390
; Sequence 390, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816

; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101014
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101068
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101071
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101471
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101472
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101474
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101479
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101915
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102207
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102240
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102307
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102330
; PRIOR FILING DATE: 1998-09-29

```

; PRIOR APPLICATION NUMBER: 60/102331
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102484
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102487
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102571
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102684
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102687
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102965
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/103258
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103314
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103315
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103328
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103395
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103401
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103633
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807

```

```

Query Match          79.2%; Score 38; DB 11; Length 146;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 NGWYPM 6
      |||||
DB      48 NGWYIM 53

```

```

RESULT 6
US-10-015-387A-390
; Sequence 390, Application US/10015387A
; Publication No. US20030135034A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC54
; CURRENT APPLICATION NUMBER: US/10/015,387A
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION removed - See File Wrapper or Palm
; SEQ ID NO 390
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-387A-390

```

```

Query Match          79.2%; Score 38; DB 12; Length 146;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 NGWYPM 6
      |||||
DB      48 NGWYIM 53

```

```

RESULT 7
US-10-006-130A-390
; Sequence 390, Application US/10006130A
; Publication No. US20030148375A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC7
; CURRENT APPLICATION NUMBER: US/10/006,130A
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 390
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-130A-390

```

```

Query Match          79.2%; Score 38; DB 12; Length 146;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
OY      1 NGWYPM 6
DB      48 NGWYIM 53

RESULT 8
US-10-199-672-376
; Sequence 376, Application US/10199672
; Publication No. US20030148442A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/199, 672
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US/10/052, 586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/06486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/06540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 376
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-672-376

Query Match      79.2%; Score 38; DB 12; Length 146;
Best Local Similarity 83.3%; Pred. NO. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 NGWYPM 6
DB      48 NGWYIM 53

RESULT 9
US-10-006-172A-390
; Sequence 390, Application US/10006172A
; Publication No. US20030153000A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C11
; CURRENT APPLICATION NUMBER: US/10/006, 172A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
```

PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102484
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102487
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102571
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102684
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102687
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102965
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 60/103258
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103314
PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: 60/103315
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103328
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103395
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103401
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103449
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103633
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103678
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103679
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103711
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/104257
PRIOR FILING DATE: 1998-10-14
PRIOR APPLICATION NUMBER: 60/104987
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105000
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105002
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105104
PRIOR FILING DATE: 1998-10-21
PRIOR APPLICATION NUMBER: 60/105169
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105266
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105693
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105694
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105807
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105881
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105882
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/106023
PRIOR FILING DATE: 1998-10-28

Query Match 79.2%; Score 38; DB 12; Length 146;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
|||||
DB 48 NGWYIM 53

RESULT 10
US-187-749-376
Sequence 376, Application US/10187749
Publication No. US2003015303641
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/187,749
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US/10/052,586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 376
LENGTH: 146
TYPE: PRT
ORGANISM: Homo Sapien
US-10-187-749-376

Query Match 79.28; Score 38; DB 12; Length 146;
Best Local Similarity 83.38; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NGWYPM 6
|||||
Db 48 NGWYIM 53

RESULT 11
US-10-194-457-376
Sequence 376, Application US/10194457
Publication No. US20030153037A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C296
CURRENT APPLICATION NUMBER: US/10/194,457
CURRENT FILING DATE: 2002-07-11
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 376
LENGTH: 146
TYPE: PRT
ORGANISM: Homo Sapien
US-10-194-457-376

Query Match 79.28; Score 38; DB 12; Length 146;
Best Local Similarity 83.38; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NGWYPM 6
|||||
Db 48 NGWYIM 53

RESULT 12
US-10-052-586-376
Sequence 376, Application US/10052586
Publication No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31

1	PRIOR FILING DATE: 1998-05-07
2	PRIOR APPLICATION NUMBER: 60/085577
3	PRIOR FILING DATE: 1998-05-15
4	PRIOR APPLICATION NUMBER: 60/085577
5	PRIOR FILING DATE: 1998-05-15
6	PRIOR APPLICATION NUMBER: 60/085580
7	PRIOR FILING DATE: 1998-05-15
8	PRIOR APPLICATION NUMBER: 60/085580
9	PRIOR FILING DATE: 1998-05-15
10	PRIOR APPLICATION NUMBER: 60/085582
11	PRIOR FILING DATE: 1998-05-15
12	PRIOR APPLICATION NUMBER: 60/085587
13	PRIOR FILING DATE: 1998-05-15
14	PRIOR APPLICATION NUMBER: 60/085700
15	PRIOR FILING DATE: 1998-05-15
16	PRIOR APPLICATION NUMBER: 60/086022
17	PRIOR FILING DATE: 1998-05-18
18	PRIOR APPLICATION NUMBER: 60/086392
19	PRIOR FILING DATE: 1998-05-22
20	PRIOR APPLICATION NUMBER: 60/086486
21	PRIOR FILING DATE: 1998-05-22
22	PRIOR APPLICATION NUMBER: 60/086486
23	PRIOR FILING DATE: 1998-05-22
24	PRIOR APPLICATION NUMBER: 60/087099
25	PRIOR FILING DATE: 1998-05-28
26	PRIOR APPLICATION NUMBER: 60/087208
27	PRIOR FILING DATE: 1998-05-28
28	PRIOR APPLICATION NUMBER: 60/087609
29	PRIOR FILING DATE: 1998-05-28
30	PRIOR APPLICATION NUMBER: 60/087759
31	PRIOR FILING DATE: 1998-06-02
32	PRIOR APPLICATION NUMBER: 60/087822
33	PRIOR FILING DATE: 1998-06-03
34	PRIOR APPLICATION NUMBER: 60/088022
35	PRIOR FILING DATE: 1998-06-04
36	PRIOR APPLICATION NUMBER: 60/088028
37	PRIOR FILING DATE: 1998-06-04
38	PRIOR APPLICATION NUMBER: 60/088028
39	PRIOR FILING DATE: 1998-06-04
40	PRIOR APPLICATION NUMBER: 60/088028
41	PRIOR FILING DATE: 1998-06-04
42	PRIOR APPLICATION NUMBER: 60/088033
43	PRIOR FILING DATE: 1998-06-04
44	PRIOR APPLICATION NUMBER: 60/088167
45	PRIOR FILING DATE: 1998-06-05
46	PRIOR APPLICATION NUMBER: 60/088202
47	PRIOR FILING DATE: 1998-06-05
48	PRIOR APPLICATION NUMBER: 60/088212
49	PRIOR FILING DATE: 1998-06-05
50	PRIOR APPLICATION NUMBER: 60/088217
51	PRIOR FILING DATE: 1998-06-05
52	PRIOR APPLICATION NUMBER: 60/088326
53	PRIOR FILING DATE: 1998-06-04
54	PRIOR APPLICATION NUMBER: 60/088655
55	PRIOR FILING DATE: 1998-06-09
56	PRIOR APPLICATION NUMBER: 60/088722
57	PRIOR FILING DATE: 1998-06-10
58	PRIOR APPLICATION NUMBER: 60/088738
59	PRIOR FILING DATE: 1998-06-10
60	PRIOR APPLICATION NUMBER: 60/088740
61	PRIOR FILING DATE: 1998-06-10
62	PRIOR APPLICATION NUMBER: 60/088811
63	PRIOR FILING DATE: 1998-06-10
64	PRIOR APPLICATION NUMBER: 60/088824
65	PRIOR FILING DATE: 1998-06-10
66	PRIOR APPLICATION NUMBER: 60/088825
67	PRIOR FILING DATE: 1998-06-10
68	PRIOR APPLICATION NUMBER: 60/088826
69	PRIOR FILING DATE: 1998-06-10
70	PRIOR APPLICATION NUMBER: 60/088861
71	PRIOR FILING DATE: 1998-06-11
72	PRIOR APPLICATION NUMBER: 60/088863
73	PRIOR FILING DATE: 1998-06-11
74	PRIOR APPLICATION NUMBER: 60/088876
75	PRIOR FILING DATE: 1998-06-11
76	PRIOR APPLICATION NUMBER: 60/089090
77	PRIOR FILING DATE: 1998-06-12
78	PRIOR APPLICATION NUMBER: 60/089105
79	PRIOR FILING DATE: 1998-06-12
80	PRIOR APPLICATION NUMBER: 60/089512
81	PRIOR FILING DATE: 1998-06-16

; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089908

Query Match 79.2%; Score 38; DB 14; Length 146;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 48 NGWYIM 53

RESULT 13

US-10-174-590-376
 ; Sequence 376, Application US/10174590
 ; Publication No. US20030008352A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C42
 ; CURRENT APPLICATION NUMBER: US/10/174,590
 ; CURRENT FILING DATE: 2002-06-18
 ; Prior Application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 376
 ; LENGTH: 146
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-174-590-376

Query Match 79.2%; Score 38; DB 15; Length 146;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 48 NGWYIM 53

RESULT 14

US-10-176-758-376
 ; Sequence 376, Application US/10176758
 ; Publication No. US20030008353A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C104
 ; CURRENT APPLICATION NUMBER: US/10/176,758
 ; CURRENT FILING DATE: 2002-06-21
 ; Prior Application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 376
 ; LENGTH: 146
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-176-758-376

Query Match 79.2%; Score 38; DB 15; Length 146;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 48 NGWYIM 53

RESULT 15

US-10-175-737-376
 ; Sequence 376, Application US/10175737
 ; Publication No. US2003003153A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C50
 ; CURRENT APPLICATION NUMBER: US/10/175,737
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior Application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 376
 ; LENGTH: 146
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-175-737-376

Query Match 79.2%; Score 38; DB 15; Length 146;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 48 NGWYIM 53

Search completed: August 20, 2003, 13:16:47
 Job time : 7.50602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:35:06 ; Search time 61.4096 Seconds
(Without alignments)
85.038 Million cell updates/sec

Title: US-09-512-082-32
Perfect score: 48
Sequence: 1 NGWYPM 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS.COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081.COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082.COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083.COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084.COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A.COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B.COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A.COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B.COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102.COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103.COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104.COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	6	14	US-09-075-338C-32
2	48	100.0	6	17	US-09-300-425B-32

3	48	100.0	6	19	US-09-512-082-32	Sequence 32, Appl
4	48	100.0	318	22	US-09-791-537-78362	Sequence 78362, A
5	43	89.6	238	27	US-10-156-761-7581	Sequence 7581, Ap
6	42	87.5	58	30	US-10-424-599-261895	Sequence 261895,
7	42	87.5	115	30	US-10-419-128-16759	Sequence 16759, A
8	42	87.5	179	30	US-10-437-963-144814	Sequence 144814,
9	42	87.5	562	1	PCT-US01-08531-33650	Sequence 33650, A
10	40	83.3	203	30	US-10-424-599-191909	Sequence 191909,
11	40	83.3	219	30	US-10-424-599-191881	Sequence 191881,
12	40	83.3	393	1	PCT-US02-28315-76	Sequence 76, Appl
13	40	83.3	403	22	US-09-799-131-92	Sequence 92, Appl
14	40	83.3	403	24	US-09-932-227-92	Sequence 92, Appl
15	40	83.3	403	27	US-10-183-708-92	Sequence 92, Appl
16	40	83.3	414	1	PCT-US00-01823-48	Sequence 48, Appl
17	40	83.3	414	18	US-09-491-577-48	Sequence 48, Appl
18	40	83.3	473	31	US-09-360-639-19292	Sequence 19292, A
19	40	83.3	667	24	US-09-902-540-10448	Sequence 10448, A
20	39	81.2	72	30	US-10-424-599-220109	Sequence 220109,
21	39	81.2	210	30	US-10-437-963-169899	Sequence 169899,
22	39	81.2	210	30	US-10-438-246-24542	Sequence 24542, A
23	39	81.2	249	16	US-09-270-767-38186	Sequence 38186, A
24	39	81.2	249	16	US-09-270-767-53403	Sequence 53403, A
25	39	81.2	249	16	US-09-270-8498-188659	Sequence 188659,
26	38	79.2	39	1	PCT-US01-00663-38311	Sequence 38311, A
27	38	79.2	39	23	US-09-864-761-42284	Sequence 42284, A
28	38	79.2	39	27	US-10-182-993-37222	Sequence 37222, A
29	38	79.2	39	27	US-10-182-995-28756	Sequence 28756, A
30	38	79.2	39	27	US-10-182-997-26552	Sequence 26552, A
31	38	79.2	39	28	US-10-203-134-38130	Sequence 38130, A
32	38	79.2	39	28	US-10-203-135-36520	Sequence 36520, A
33	38	79.2	39	28	US-10-203-136-38125	Sequence 38125, A
34	38	79.2	39	28	US-10-203-137-38311	Sequence 38311, A
35	38	79.2	39	28	US-10-203-139-36729	Sequence 36729, A
36	38	79.2	49	30	US-10-437-963-134023	Sequence 134023,
37	38	79.2	75	30	US-10-424-599-264214	Sequence 264214,
38	38	79.2	81	19	US-09-513-996A-60571	Sequence 60571, A
39	38	79.2	83	19	US-09-513-996A-4640	Sequence 4640, Ap
40	38	79.2	99	1	PCT-US02-32727-8826	Sequence 8826, Ap
41	38	79.2	99	25	US-09-978-825-8826	Sequence 8826, Ap
42	38	79.2	99	26	US-10-057-498-8826	Sequence 8826, Ap
43	38	79.2	146	1	PCT-US00-26524B-5160	Sequence 5160, Ap
44	38	79.2	146	1	PCT-US03-11497-93	Sequence 93, Appl
45	38	79.2	146	24	US-09-946-374-390	Sequence 390, App

ALIGNMENTS

RESULT 1
US-09-075-338C-32
Sequence 32, Application US/09075338C
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARLI, Lorenzo
APPLICANT: VIRI, Francesca
APPLICANT: BIRCHLER, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
FILE REFERENCE: SCH-1733
CURRENT APPLICATION NUMBER: US/09/075,338C
CURRENT FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
OTHER INFORMATION: antibody clone
US-09-075-338C-32

Query Match 100.0%; Score 48; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NGWYPM 6
|||||
Db 1 NGWYPM 6

RESULT 2
US-09-300-425B-32
; Sequence 32, Application US/09300425B
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-173P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-300-425B-32

Query Match 100.0%; Score 48; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6
|||||
Db 1 NGWYPM 6

RESULT 3
US-09-512-082-32
; Sequence 32, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-173P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-512-082-32

Query Match 100.0%; Score 48; DB 19; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NGWYPM 6
|||||
Db 1 NGWYPM 6

RESULT 4
US-09-791-537-78362
; Sequence 78362, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 78362
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-78362

Query Match 100.0%; Score 48; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6
|||||
Db 222 NGWYPM 227

RESULT 5
US-10-156-761-7581
; Sequence 7581, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARDO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 7581
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7581

Query Match 89.6%; Score 43; DB 27; Length 318;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6
|||||
Db 106 SGWYPM 111

RESULT 6

US-10-424-599-261895
; Sequence 261895, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 261895
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78513C.1.pep
US-10-424-599-261895

Query Match 87.5%; Score 42; DB 30; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
DB 35 GWYPM 39

RESULT 7
US-10-419-128-16759
; Sequence 16759, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16759
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-16759

Query Match 87.5%; Score 42; DB 30; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
DB 56 GWYPM 60

RESULT 8
US-10-437-963-144814
; Sequence 144814, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad

APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144814
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45594C.1.pep
US-10-437-963-144814

Query Match 87.5%; Score 42; DB 30; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
DB 116 GWYPM 120

RESULT 9
PCT-US01-08631-33650
; Sequence 33650, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 33650
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-33650

Query Match 87.5%; Score 42; DB 1; Length 562;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
DB 367 GWYPM 371

RESULT 10
US-10-424-599-191909
; Sequence 191909, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 191909
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_15131C.1.pcp
US-10-424-599-191909

Query Match 83.3%; Score 40; DB 30; Length 203;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 158 NGWYPM 163

RESULT 11

Sequence 191881, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 191881
LENGTH: 219
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(219)
OTHER INFORMATION: unsure at all xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_15289C.1.pcp
US-10-424-599-191881

Query Match 83.3%; Score 40; DB 30; Length 219;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 148 NGWYPM 153

RESULT 12

PCT-US02-28315-76
Sequence 76, Application PC/TUS0228315
GENERAL INFORMATION:
APPLICANT: Sentigen Corp.
APPLICANT: Lee, Kevin
APPLICANT: Ong, Jane
APPLICANT: Nguyen, Thuy-Ai
APPLICANT: Kloss, Brian
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF INSECT O-83B ODORANT RECEPTOR
TITLE OF INVENTION: GENES AND USES THEREOF
FILE REFERENCE: 10657-005-228
CURRENT APPLICATION NUMBER: PCT/US02/28315
CURRENT FILING DATE: 2002-09-04
PRIOR APPLICATION NUMBER: 60/317,407
PRIOR FILING DATE: 2001-09-04
NUMBER OF SEQ ID NOS: 97
SOFTWARE: PatentIn version 3.0
SEQ ID NO 76
LENGTH: 393
TYPE: PRT
ORGANISM: Anopheles gambiae
PCT-US02-28315-76

Query Match 83.3%; Score 40; DB 1; Length 393;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 176 NTWYPM 181

RESULT 13

US-09-799-131-92
Sequence 92, Application US/09799131
GENERAL INFORMATION:
APPLICANT: VOSSHALL, LESLIE
APPLICANT: AMREIN, HUBERT
APPLICANT: AXEL, RICHARD
TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
FILE REFERENCE: 0575/58715-B/JPM/ADM/BJA
CURRENT APPLICATION NUMBER: US/09/799,131
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: PCT/US00/04995
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: PatentIn version 3.0
SEQ ID NO 92
LENGTH: 403
TYPE: PRT
ORGANISM: Drosophila Melanogaster
US-09-799-131-92

Query Match 83.3%; Score 40; DB 22; Length 403;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 176 NTWYPM 181

RESULT 14

US-09-932-227-92
Sequence 92, Application US/09932227
GENERAL INFORMATION:
APPLICANT: VOSSHALL, LESLIE
APPLICANT: AMREIN, HUBERT
APPLICANT: AXEL, RICHARD
TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
FILE REFERENCE: 0575/58715-A-PCT-US/JPM/ADM/BJA
CURRENT APPLICATION NUMBER: US/09/932,227
CURRENT FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: PCT/US00/04995
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn version 3.0
SEQ ID NO 92
LENGTH: 403
TYPE: PRT
ORGANISM: Drosophila Melanogaster
US-09-932-227-92

Query Match 83.3%; Score 40; DB 24; Length 403;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
DB 176 NTWYPM 181

RESULT 15

US-10-183-708-92
Sequence 92, Application US/10183708
GENERAL INFORMATION:
APPLICANT: VOSSHALL, LESLIE
APPLICANT: AMREIN, HUBERT

; APPLICANT: AXEL, RICHARD
; TITLE OF INVENTION: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF
; FILE REFERENCE: 0575/58715-AA-PCT-US/JPM/ADM/BJA
; CURRENT APPLICATION NUMBER: US/10/183,708
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 09/932,227
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/04995
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/257,706
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Drosophila Melanogaster DOR83
US-10-183-708-92

Query Match 83.3%; Score 40; DB 27; Length 403;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
1 1111
Db 176 NTWYPM 181

Search completed: August 20, 2003, 13:13:43
Job time : 62.4096 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 1.73494 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-32

Perfect score: 48

Sequence: 1 NGWYPM 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US00_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	37.5	78.1	542	1 PCT-US03-10753-40	Sequence 40, Appl
2	37.5	78.1	552	1 PCT-US03-10753-37	Sequence 37, Appl
3	37	77.1	512	6 US-10-408-765A-2774	Sequence 2774, Ap
4	36	75.0	258	6 US-10-408-765A-878	Sequence 878, App
5	36	75.0	258	6 US-10-408-765A-1572	Sequence 1572, Ap
6	36	75.0	330	6 US-10-603-114-7860	Sequence 7860, Ap
7	36	75.0	401	1 PCT-US03-07174-20	Sequence 20, Appl
8	36	75.0	472	6 US-10-612-783-6923	Sequence 6923, Ap
9	36	75.0	481	6 US-10-612-783-6924	Sequence 6924, Ap
10	36	75.0	634	6 US-10-286-897-2453	Sequence 2453, Ap
11	36	75.0	634	6 US-10-258-898A-2453	Sequence 2453, Ap
12	36	75.0	1063	6 US-10-450-727-1	Sequence 1, Appl
13	36	75.0	1063	6 US-10-408-765A-2458	Sequence 2458, Ap
14	36	75.0	1291	6 US-10-286-897-6025	Sequence 6025, Ap
15	36	75.0	1291	6 US-10-258-898A-6025	Sequence 6025, Ap
16	35	72.9	273	6 US-10-371-525A-30	Sequence 30, Appl
17	35	72.9	273	6 US-10-371-645-30	Sequence 30, Appl
18	35	72.9	273	6 US-10-371-260-30	Sequence 30, Appl
19	35	72.9	419	6 US-10-603-114-6107	Sequence 6107, Ap
20	34	70.8	111	6 US-10-273-573-8697	Sequence 8697, Ap
21	34	70.8	186	6 US-10-293-244-1550	Sequence 1550, Ap
22	34	70.8	210	6 US-10-293-244-3518	Sequence 3518, Ap
23	34	70.8	230	6 US-10-627-476-604	Sequence 604, App
24	34	70.8	238	6 US-10-603-113-14629	Sequence 14629, A
25	34	70.8	248	6 US-10-603-114-5430	Sequence 5430, Ap
26	34	70.8	284	6 US-10-466-531-122	Sequence 122, App

27	34	70.8	307	6 US-10-292-798-1442	Sequence 1442, Ap
28	34	70.8	312	6 US-10-637-011-10	Sequence 10, Appl
29	34	70.8	312	6 US-10-637-011-165	Sequence 165, App
30	34	70.8	312	6 US-10-637-011-166	Sequence 166, App
31	34	70.8	314	6 US-10-637-011-24	Sequence 24, Appl
32	34	70.8	502	6 US-10-273-573-7981	Sequence 7981, Ap
33	34	70.8	542	6 US-10-637-011-17	Sequence 17, Appl
34	34	70.8	566	6 US-10-637-011-18	Sequence 18, Appl
35	34	70.8	577	6 US-10-637-011-3	Sequence 3, Appl
36	34	70.8	686	6 US-10-408-765A-1545	Sequence 1545, Ap
37	34	70.8	803	6 US-10-603-114-5399	Sequence 5399, Ap
38	34	70.8	1139	7 US-60-487-610-1772	Sequence 1772, Ap
39	34	70.8	1139	7 US-60-485-450-1138	Sequence 1138, Ap
40	34	70.8	1170	6 US-10-089-3208-100	Sequence 100, App
41	34	70.8	1173	7 US-60-487-610-1771	Sequence 1771, Ap
42	34	70.8	1173	7 US-60-485-450-1137	Sequence 1137, Ap
43	33.5	69.8	569	1 PCT-US03-10753-39	Sequence 39, Appl
44	33	68.8	25	6 US-10-627-631-29	Sequence 29, Appl
45	33	68.8	59	6 US-10-273-573-6716	Sequence 6716, Ap

ALIGNMENTS

RESULT 1
PCT-US03-10753-40
Sequence 40, Application PC/TUS0310753
GENERAL INFORMATION:
APPLICANT: Schockey, Jay M.
APPLICANT: Schmitt, Judy
TITLE OF INVENTION: Plant Acyl-CoA Synthetases
FILE REFERENCE: DOM-07655
CURRENT APPLICATION NUMBER: PCT/US03/10753
CURRENT FILING DATE: 2003-04-09
PRIOR APPLICATION NUMBER: 10/119,136
PRIOR FILING DATE: 2002-04-09
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 187
SOFTWARE: PatentIn version 3.2
SEQ ID NO 40
LENGTH: 542
TYPE: PRT
ORGANISM: Arabidopsis thaliana
PCT-US03-10753-40

Query Match 78.1% Score 37.5; DB 1; Length 542;
Best Local Similarity 85.7% Pred. No. 1e+02; 0; Indels 1; Gaps 1;
Matches 6; Conservative 0; Mismatches 0;

Qy 1 NGWYPM 6
Db 236 NGWYPM 242

RESULT 2
PCT-US03-10753-37
Sequence 37, Application PC/TUS0310753
GENERAL INFORMATION:
APPLICANT: Schockey, Jay M.
APPLICANT: Schmitt, Judy
TITLE OF INVENTION: Plant Acyl-CoA Synthetases
FILE REFERENCE: DOM-07655
CURRENT APPLICATION NUMBER: PCT/US03/10753
CURRENT FILING DATE: 2003-04-09
PRIOR APPLICATION NUMBER: 10/119,136
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 10/410,031
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 187
SOFTWARE: PatentIn version 3.2

SEQ ID NO 37
LENGTH: 552
TYPE: PRT
ORGANISM: Arabidopsis thaliana
PCT-US03-10753-37

Query Match 78.1%; Score 37.5; DB 1; Length 552;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 NGW-YWP 6
DB 239 NGWSTPW 245

RESULT 3
US-10-408-765A-2774
; Sequence 2774, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2774
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2774

Query Match 77.1%; Score 37; DB 6; Length 512;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GYWPW 6
DB 244 GWHWP 248

RESULT 4
US-10-408-765A-878
; Sequence 878, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 878
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-878

Query Match 75.0%; Score 36; DB 6; Length 258;

Best Local Similarity 100.0%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYWP 6
DB 199 WYWP 202

RESULT 5
US-10-408-765A-1572
; Sequence 1572, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1572
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1572

Query Match 75.0%; Score 36; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYWP 6
DB 199 WYWP 202

RESULT 6
US-10-603-114-7860
; Sequence 7860, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRA
; FILE REFERENCE: 2709.1002-001
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7860
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-7860

Query Match 75.0%; Score 36; DB 6; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYWP 6
DB 5 WYWP 8

RESULT 7
PCT-US03-07174-20

Sequence 20, Application PC/TUS0307174
GENERAL INFORMATION:
APPLICANT: VANDERBILT UNIVERSITY
TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
FILE REFERENCE: N9267
CURRENT APPLICATION NUMBER: PCT/US03/07174
CURRENT FILING DATE: 2003-05-12
PRIOR APPLICATION NUMBER: 10/094,240
PRIOR FILING DATE: 2002-03-08
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 401
TYPE: PRT
ORGANISM: Anopheles gambiae
PCT-US03-07174-20

Query Match
Best Local Similarity 75.0%; Score 36; DB 1; Length 401;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
DB 133 WYPW 136

RESULT 8
US-10-612-783-6923
Sequence 6923, Application US/10612783
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(53373)A
CURRENT APPLICATION NUMBER: US/10/612,783
CURRENT FILING DATE: 2003-07-02
NUMBER OF SEQ ID NOS: 7098
SEQ ID NO 6923
LENGTH: 472
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_6952C.1.pep
US-10-612-783-6923

Query Match
Best Local Similarity 75.0%; Score 36; DB 6; Length 472;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
DB 417 WYPW 420

RESULT 9
US-10-612-783-4924
Sequence 4924, Application US/10612783
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(53373)A
CURRENT APPLICATION NUMBER: US/10/612,783
CURRENT FILING DATE: 2003-07-02
NUMBER OF SEQ ID NOS: 7098
SEQ ID NO 4924
LENGTH: 481

TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(481)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4577_160568C.1.pep
US-10-612-783-4924

Query Match
Best Local Similarity 75.0%; Score 36; DB 6; Length 481;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
DB 426 WYPW 429

RESULT 10
US-10-286-897-2453
Sequence 2453, Application US/10286897
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/10/286,897
CURRENT FILING DATE: 2002-11-01
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US/09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US/09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US/09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US/09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US/09/662,191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US/09/693,036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US/09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7143
SOFTWARE: pt_FL_genes_b Versions 1.0
SEQ ID NO 2453
LENGTH: 634
TYPE: PRT
ORGANISM: Homo sapiens
US-10-286-897-2453

Query Match
Best Local Similarity 75.0%; Score 36; DB 6; Length 634;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
DB 309 WYPW 312

RESULT 11
US-10-258-898A-2453
Sequence 2453, Application US/10258898A
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/10/258,898A
CURRENT FILING DATE: 2002-10-29
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317

```
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: US09/598,042
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: US09/620,312
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: US09/653,450
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: US09/662,191
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: US09/693,036
;; PRIOR FILING DATE: 2000-10-19
;; PRIOR APPLICATION NUMBER: US09/727,344
;; PRIOR FILING DATE: 2000-11-29
;; NUMBER OF SEQ ID NOS: 7143
;; SOFTWARE: pt.fl_genes_b Versions 1.0
;; SEQ ID NO: 2453
;; LENGTH: 634
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-258-898A-2453
```

```
Query Match      75.0%; Score 36; DB 6; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 WYPM 6
        ||||
Db      309 WYPM 312
```

```
RESULT 12
US-10-450-727-1
;; Sequence 1, Application US/10450727
;; GENERAL INFORMATION:
;; APPLICANT: INCYTE GENOMICS, INC.
;; APPLICANT: LEE, Ernestine A.
;; APPLICANT: BAUGHN, Mariah R.
;; TITLE OF INVENTION: AMINOACYL TRNA SYNTHETASES
;; FILE REFERENCE: PI-0323 PCT
;; CURRENT APPLICATION NUMBER: US/10/450,727
;; CURRENT FILING DATE: 2003-06-13
;; PRIOR APPLICATION NUMBER: 60/255,963
;; PRIOR FILING DATE: 2000-12-15
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: PERL Program
;; SEQ ID NO: 1
;; LENGTH: 1063
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: incyte ID No: 60211391CD1
US-10-450-727-1
```

```
Query Match      75.0%; Score 36; DB 6; Length 1063;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 WYPM 6
        ||||
Db      113 WYPM 116
```

```
RESULT 13
US-10-408-765A-2458
;; Sequence 2458, Application US/10408765A
;; GENERAL INFORMATION:
;; APPLICANT: Ghosh, Soumitra S.
;; APPLICANT: Fahy, Eoin D.
;; APPLICANT: Zhang, Bing
;; APPLICANT: Gibson, Bradford W.
;; APPLICANT: Taylor, Steven W.
;; APPLICANT: Glenn, Gary M.
```

```
;; APPLICANT: Warnock, Dale E.
;; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;; FILE REFERENCE: 660088.465
;; CURRENT APPLICATION NUMBER: US/10/408,765A
;; CURRENT FILING DATE: 2003-04-04
;; NUMBER OF SEQ ID NOS: 3077
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 2458
;; LENGTH: 1063
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-408-765A-2458
```

```
Query Match      75.0%; Score 36; DB 6; Length 1063;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 WYPM 6
        ||||
Db      113 WYPM 116
```

```
RESULT 14
US-10-286-897-6025
;; Sequence 6025, Application US/10286897
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq Inc
;; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;; FILE REFERENCE: 784FLPCT
;; CURRENT APPLICATION NUMBER: US/10/286,897
;; CURRENT FILING DATE: 2002-11-01
;; PRIOR APPLICATION NUMBER: US/09/488,725
;; PRIOR FILING DATE: 2000-01-21
;; PRIOR APPLICATION NUMBER: US/09/552,317
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: US/09/598,042
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: US/09/620,312
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: US/09/653,450
;; PRIOR FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: US/09/662,191
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: US/09/693,036
;; PRIOR FILING DATE: 2000-10-19
;; PRIOR APPLICATION NUMBER: US/09/727,344
;; PRIOR FILING DATE: 2000-11-29
;; NUMBER OF SEQ ID NOS: 7143
;; SOFTWARE: pt.fl_genes_b Versions 1.0
;; SEQ ID NO: 6025
;; LENGTH: 1291
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-286-897-6025
```

```
Query Match      75.0%; Score 36; DB 6; Length 1291;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      3 WYPM 6
        ||||
Db      332 WYPM 335
```

```
RESULT 15
US-10-258-898A-6025
;; Sequence 6025, Application US/10258898A
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq Inc
;; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;; FILE REFERENCE: 784FLPCT
;; CURRENT APPLICATION NUMBER: US/10/258,898A
```

; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_fl_genes_b Versions 1.0
; SEQ ID NO 6025
; LENGTH: 1291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-6025

Query Match 75.0%; Score 36; DB 6; Length 1291;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
Db 332 WYPW 335

Search completed: August 20, 2003, 12:45:14
Job time: 1.73494 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 4.19277 Seconds
(without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-32

Perfect score: 48
Sequence: 1 NGWYPW 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	87.5	126	2 H64236	hypothetical prote
2	42	87.5	197	2 S73688	hypothetical prote
3	42	87.5	198	2 H90550	hypothetical prote
4	42	87.5	336	2 S72858	hypothetical prote
5	40	83.3	456	2 T51117	hypothetical prote
6	39	81.2	371	2 S55490	hypothetical prote
7	39	81.2	708	2 T22377	hypothetical prote
8	38	79.2	223	2 A86249	protein T23018.24
9	38	79.2	376	2 T40488	hypothetical prote
10	38	79.2	764	2 E90257	hypothetical prote
11	37.5	78.1	540	2 B83121	probable AMP-bind
12	37.5	78.1	900	2 JH0157	cellulase (EC 3.2.
13	37	77.1	250	2 S43104	outer membrane pro
14	37	77.1	262	2 T36541	hypothetical prote
15	37	77.1	269	2 A44611	outer membrane pro
16	37	77.1	272	2 C81893	outer membrane pro
17	37	77.1	272	2 A81126	class 5 outer memb
18	37	77.1	352	2 S77448	hypothetical prote
19	37	77.1	1559	1 S64757	probable membrane
20	37	77.1	1797	2 T21889	hypothetical prote
21	37	77.1	1805	2 T21888	hypothetical prote
22	36	75.0	190	2 T21888	hypothetical prote
23	36	75.0	210	2 B90082	hypothetical prote
24	36	75.0	210	2 B90082	hypothetical prote
25	36	75.0	210	2 B90118	hypothetical prote
26	36	75.0	210	2 B90118	hypothetical prote
27	36	75.0	210	2 B90138	hypothetical prote
28	36	75.0	210	2 F84211	hypothetical prote
29	36	75.0	214	2 A64483	hypothetical prote

30	36	75.0	214	2 A64302	transposase - Meth
31	36	75.0	236	2 JQ0606	arylesterase (EC 3
32	36	75.0	334	2 S35523	DNA-directed DNA p
33	36	75.0	334	2 E90813	DNA polymerase III
34	36	75.0	334	2 A85673	DNA polymerase III
35	36	75.0	334	2 AB0643	DNA polymerase III
36	36	75.0	340	2 A10195	DNA-directed DNA p
37	36	75.0	345	2 F69200	conserved hypotet
38	36	75.0	349	2 H72714	probable O-sialogl
39	36	75.0	366	2 S53311	hydroxymandelonitr
40	36	75.0	423	1 A29639	carboxypeptidase D
41	36	75.0	425	2 E84631	probable serine ca
42	36	75.0	447	2 G84772	probable serine ca
43	36	75.0	452	2 H84772	probable serine ca
44	36	75.0	465	2 B85358	SERINE CARBOXYPEPT
45	36	75.0	474	2 D84631	probable serine ca

ALIGNMENTS

RESULT 1

H64236
hypothetical protein MG333 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C:Accession: H64236
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.M.; Fumrmann, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; NCID:96026346; PMID:7569993
A:Accession: H64236
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-126 <TIGR>
A:Cross-references: GB:U39715; GB:I43967; NID:g1046026; PID:g1046036; TIGR:MG333
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: acyl carrier protein phosphodiesterase

Query Match 87.5%; Score 42; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPW 6
Db 68 GWYPW 72

RESULT 2
S73688
hypothetical protein H11366 - Mycoplasma pneumoniae (strain ATCC 29342)
N:Alternate names: hypothetical protein P01_orf197
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 17-Mar-2000
C:Accession: S73688
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996.
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon
A:Reference number: S73327; NCID:97105885; PMID:8948633
A:Accession: S73688
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-197 <HIM>
A:Cross-references: EMBL:AE000034; GB:U00089; NID:g1674031; PIDN:AA96010.1; PID:g167
C:Genetics:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetic code: SGC3
C:Superfamily: acyl carrier protein phosphodiesterase

Query Match 87.5%; Score 42; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYWP 6
 |||||
 DB 139 GWYWP 143

RESULT 3

H90550

hypothetical protein MYPV_3120 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: H90550
 R:Chambaud, I.; Heillig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
 A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: H90550

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-198 <KUR>

A:Cross-references: GB:AL445566; PID:q14089726; PIDN:CAC13485.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV_3120

A:Genetic code: SGC3

Query Match 87.5%; Score 42; DB 2; Length 198;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYWP 6
 |||||
 DB 139 GWYWP 143

RESULT 4

S72858

hypothetical protein B2126_C2_219 - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001

C:Accession: S72858

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B2126.

A:Reference number: S72585

A:Accession: S72858

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-336 <SMI>

A:Cross-references: EMBL:U00017; NID:9466994; PIDN:AA17198.1; PID:9467013

Query Match 87.5%; Score 42; DB 2; Length 336;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYWP 6
 |||||
 DB 3 GWYWP 7

RESULT 5

T51117

hypothetical protein [imported] - Brevibacterium linens (fragment)

C:Species: Brevibacterium linens

C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000

C:Accession: T51117

R:Kruusik, P.; Sandmann, G.

Mol. Genet. 263, 423-432, 2000

A:Title: A carotenogenic gene cluster from Brevibacterium linens with novel lycopene cyc

A:Reference number: Z25303; MUID:20279196; PMID:10821176
 A:Accession: T51117
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-456 <KRJ>
 A:Cross-references: EMBL:AF139916; PIDN:AAF65580.1
 A:Experimental source: DSM 20426; ATCC9175

Query Match 83.3%; Score 40; DB 2; Length 456;
 Best Local Similarity 83.3%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYWP 6
 |||||
 DB 333 NGWYWP 338

RESULT 6

S55490

hypothetical protein SPAC5H10.12c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 31-Jan-2000

C:Accession: T38976; S55490

R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, May 1995

A:Reference number: Z21821

A:Accession: T38976

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-371 <CO2>

A:Cross-references: EMBL:Z49811; NID:9854599; PIDN:CAA89962.1; PID:9854611; GSPDB:GNO

A:Experimental source: strain 972h; cosmid c5H10

C:Genetics:

A:Gene: SPAC5H10.12c

A:Map position: 1

Query Match 81.2%; Score 39; DB 2; Length 371;
 Best Local Similarity 66.7%; Pred. No. 70;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYWP 6
 |||||
 DB 307 NGWYWP 312

RESULT 7

T22377

hypothetical protein F48C11.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22377

R:Wild, A.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19556

A:Accession: T22377

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-708 <WIL>

A:Cross-references: EMBL:Z80789; PIDN:CAR02551.1; GSPDB:GN00028; CESP:F48C11.2

A:Experimental source: clone F48C11

C:Genetics:

A:Gene: CESP:F48C11.2

A:Map position: X

Query Match 81.2%; Score 39; DB 2; Length 708;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYWP 6
 |||||
 DB 696 NGWYWP 701

RESULT 8
A86249
protein T23j18.24 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: A86249
R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86249
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-223 <STO>
A:Cross-references: GB:AE005172; NID:96554191; PIDN:AAFI6637.1; GSPDB:GN00141
C:Genetics:
A:Gene: T23j18.24
A:Map position: 1
C:Superfamily: pectinesterase

Query Match 79.2%; Score 38; DB 2; Length 223;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GMYPW 6
11:11
DB 152 GMYPW 156

RESULT 9
T40488
hypothetical protein SPBC4C3.09 duplicated in pombe - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40488
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21910
A:Accession: T40488
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-376 <WOO>
A:Cross-references: EMBL:AL021730; PIDN:CAA16831.1; GSPDB:GN00067; SPDB:SPBC4C3.09
C:Genetics:
A:Experimental source: strain 972h; cosmid c4C3
A:Gene: SPDB:SPBC4C3.09
A:Map position: 2

Query Match 79.2%; Score 38; DB 2; Length 376;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GMYPW 6
11:11
DB 313 GMYPW 317

RESULT 10
E90257
hypothetical protein SEO1052 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: E90257
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.

arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-764 <KOR>
A:Cross-references: GB:AE006641; NID:913814238; PIDN:AAK1316.1; GSPDB:GN00155
C:Genetics:
A:Gene: SEO1052

Query Match 79.2%; Score 38; DB 2; Length 764;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GMYPW 6
11:11
DB 74 GMYPW 78

RESULT 11
B83121
Probable AMP-binding enzyme PA4198 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83121
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.
Loay, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83121
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-540 <STO>
A:Cross-references: GB:AE004836; GB:AE004091; NID:99950405; PIDN:AAG07585.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4198

Query Match 78.1%; Score 37.5; DB 2; Length 540;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 NGMYPW 6
111111
DB 235 NGMYPW 241

RESULT 12
JH0157
cellulase (EC 3.2.1.4) H precursor - Clostridium thermocellum
N:Alternate names: endo-1,4-beta-glucanase; endoglucanase H
C:Species: Clostridium thermocellum
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 15-Oct-1999
C:Accession: JH0157
R:Yaguee, E.; Beguin, P.; Aubert, J.P.
Gene 89, 61-67, 1990.
A:Title: Nucleotide sequence and deletion analysis of the cellulase-encoding gene cel
A:Reference number: JH0157; MUID:90323606; PMID:2197182
A:Accession: JH0157
A:Molecule type: DNA
A:Residues: 1-900 <YAG>
A:Cross-references: GB:M31903; NID:9144773; PIDN:AA23225.1; PID:9144774
A:Note: The authors translated the codon CAG for residue 863 as His
C:Comment: Cellulase H is involved in the hydrolysis of cellulose and arranged in cel
C:Comment: Cellulase H hydrolyzes carboxymethylcellulose, p-nitrophenyl-beta-D-cellob
C:Genetics:
A:Gene: celH
A:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as

A:Pathway: cellulose degradation
 C:Superfamily: Clostridium cellulase repeat homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-44/Domain: signal sequence #status predicted <SIG>
 F:45-900/Product: cellulase H #status predicted <CEL>
 F:309-324/Region: proline/serine/threonine-rich
 F:631-654/Region: proline/serine/threonine-rich
 F:833-856/Domain: Clostridium cellulase repeat homology <CCRI>
 F:872-895/Domain: Clostridium cellulase repeat homology <CCR2>

Query Match 78.1%; Score 37.5; DB 2; Length 900;
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 NGWY--PW 6
 |||||
 DB 133 NGDWYFW 139

RESULT 13

S43104
 outer membrane protein class 5c - Neisseria meningitidis (fragment)
 C:Species: Neisseria meningitidis
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
 C:Accession: S43104
 R:Guillen, G.; Leal, M.J.; Alvarez, A.; Delgado, M.; Silva, R.; Herrera, L.
 submitted to the EMBL Data Library, March 1994
 A:Description: Cloning and expression of the Neisseria meningitidis 5C outer membrane pr
 A:Reference number: S43104
 A:Accession: S43104
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-250 <GDI>
 A:Cross-references: EMBL:X78221; NID:g467599; PIDN:CA55065.1; PID:g984673

Query Match 77.1%; Score 37; DB 2; Length 250;
 Best Local Similarity 75.0%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

OY 1 NGWY--PW 6
 |||||
 DB 151 NGWYINPW 158

RESULT 14

T36541
 hypothetical protein SCH1029c1 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T36541
 R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, May 1999
 A:Reference number: Z21609
 A:Accession: T36541
 A>Status: preliminary; translated from GB/EMBL/DDBT
 A:Molecule type: DNA
 A:Residues: 1-262 <MUR>
 A:Cross-references: EMBL:AL049754; PIDN:CAB42038.1; GSPDB:GNO0070; SCOEDB:SCH10.29C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCH10.29C

Query Match 77.1%; Score 37; DB 2; Length 262;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYFW 6
 ||:|
 DB 141 GWHFW 145

RESULT 15

A44611

outer membrane protein class 5c precursor - Neisseria meningitidis (strain Z3476)

C:Species: Neisseria meningitidis
 C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
 C:Accession: A44611
 R:Olyhoek, A.J.M.; Sarkari, J.; Bopp, M.; Morelli, G.; Achtman, M.
 Microb. Pathog. 11, 249-257, 1991
 A>Title: Cloning and expression in Escherichia coli of opc, the gene for an unusual c
 A:Reference number: A44611; MUID:92261288; PMID:1813777
 A:Accession: A44611
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <OLY>
 A:Cross-references: GB:M80195
 C:Genetics:
 A:Gene: opc

Query Match 77.1%; Score 37; DB 2; Length 269;
 Best Local Similarity 75.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

OY 1 NGWY--PW 6
 |||||
 DB 170 NGWYINPW 177

Search completed: August 20, 2003, 12:42:24
 Job time: 5.19277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 2.27711 Seconds

(Without alignments) 123.912 Million cell updates/sec

Title: US-09-512-082-32

Perfect score: 48

Sequence: 1 NGWYPPW 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues 127863

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt.41.*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	87.5	196	1 ACPD_MYCGE	P47575 mycoplasma
2	42	87.5	197	1 ACPD_MYCPN	P75305 mycoplasma
3	42	87.5	198	1 ACPD_MYCPN	P75305 mycoplasma
4	40	83.3	393	1 069A_DROME	P82985 drosophila
5	39	81.2	371	1 YAOC_SCHPO	P09680 schizosacch
6	38	79.2	319	1 PME_DAUCA	P83218 daucus caro
7	37.5	78.1	900	1 GUNH_CLOTM	P14772 clostridium
8	37	77.1	1559	1 BPT1_YEAST	P60323 methanococ
9	36	75.0	214	1 Y017_MERJA	P55747 hordeum vul
10	36	75.0	214	1 Y017_MERJA	P55747 hordeum vul
11	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
12	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
13	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
14	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
15	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
16	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
17	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
18	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
19	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
20	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
21	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
22	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
23	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
24	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
25	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
26	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
27	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
28	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
29	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
30	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
31	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
32	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul
33	36	75.0	254	1 Y017_MERJA	P55747 hordeum vul

34	35	72.9	2195	1 SC16_YEAST	P48415 saccharomyc
35	34	70.8	167	1 REV_VITIV	P21280 viana lenti
36	34	70.8	167	1 REV_VITIV	P21280 viana lenti
37	34	70.8	167	1 REV_VITIV	P21280 viana lenti
38	34	70.8	167	1 REV_VITIV	P21280 viana lenti
39	34	70.8	167	1 REV_VITIV	P21280 viana lenti
40	34	70.8	167	1 REV_VITIV	P21280 viana lenti
41	34	70.8	167	1 REV_VITIV	P21280 viana lenti
42	34	70.8	167	1 REV_VITIV	P21280 viana lenti
43	34	70.8	167	1 REV_VITIV	P21280 viana lenti
44	34	70.8	167	1 REV_VITIV	P21280 viana lenti
45	34	70.8	167	1 REV_VITIV	P21280 viana lenti

ALIGNMENTS

RESULT 1
ACPD_MYCGE STANDARD; PRT; 196 AA.
ID ACPD_MYCGE
AC P47575; 049357;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP
phosphodiesterase).
GN MG333.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=96026346; Pubmed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
Romb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
"The minimal gene complement of Mycoplasma genitalium";
Science 270:397-403(1995).
[2]
RP SEQUENCE OF 1-47 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=94075230; Pubmed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
"A survey of the Mycoplasma genitalium genome by using random
sequencing";
J. Bacteriol. 175:7918-7930(1993).
[3]
RP CONCEPTUAL TRANSLATION.
RA Balroch A.;
RL Unpublished observations (FEB-1997).
CC -I- FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of
the phosphopantetheine prosthetic group from ACP (by similarity).
CC -I- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-
phosphopantetheine + apo-[acyl-carrier protein].
CC -I- SIMILARITY: Belongs to the acpI family.
CC -I- CAUTION: THIS IS A CONCEPTUAL TRANSLATION, A FRAMESHIFT WAS
INTRODUCED IN POSITION 74 TO MAXIMIZE THE SIMILARITY WITH OTHER
ACPD.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb.ch).
CC EMBL: U39714; -; NOT_ANNOTATED_CDS.
CC EMBL: U02249; AADI2538.1; ALT_INIT.

TIGR: MG333; -. 1.
 DR HAMAP; MF_01216; -. 1.
 DR InterPro; IPR003680; NADHdh_2.
 DR Pfam; PF02525; Flavodoxin_2; 1.
 DR Hypothetical protein; Hydrolase; Complete proteome.
 FT CONFLICT 44 47 NGLP -> MNYQ (IN REF. 3).
 KW MYPV_3120.
 OS Mycoplasma pulmonis.
 SEQUENCE 196 AA; 22076 MW; 9D26C167A3DBC88C CRC64;

Query Match 87.5%; Score 42; DB 1; Length 196;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWYPM 6
 |||||
 DB 139 GWYPM 142

RESULT 2
 ACPD_MYCPN STANDARD; PRT; 197 AA.
 AC P75305; .
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP phosphodiesterase).
 GN MPM479 OR MP362.
 OS Mycoplasma pneumoniae.
 CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OC NCBI_TaxId=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RA MEDLINE=97105885; PubMed=8948633;
 RA Himmelfeich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
 RA Hermann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";
 RT Nucleic Acids Res. 24:4420-4449(1996).
 RL FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of the phosphopantetheine prosthetic group from ACP (By similarity).
 CC -1- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-phosphopantetheine + apo-[acyl-carrier protein].
 CC -1- SIMILARITY: Belongs to the acpd family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000034; AAB96010.1; -.
 DR PIR; S73688; S73688.
 DR HAMAP; MF_01216; -. 1.
 DR InterPro; IPR003680; NADHdh_2.
 DR Pfam; PF02525; Flavodoxin_2; 1.
 DR Hypothetical protein; Hydrolase; Complete proteome.
 KW SEQUENCE 197 AA; 21550 MW; B566B394F793631 CRC64;

Query Match 87.5%; Score 42; DB 1; Length 197;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWYPM 6
 |||||
 DB 139 GWYPM 143

RESULT 3
 ACPD_MYCPN STANDARD; PRT; 198 AA.
 ID ACPD_MYCPN

Q980P9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP phosphodiesterase).
 GN MYPV_3120.
 OS Mycoplasma pulmonis.
 CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OC NCBI_TaxId=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAB CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chamblaud I., Hellig R., Ferris S., Barbe V., Samson D., Galisson F., Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C., Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis.";
 RT Nucleic Acids Res. 29:2145-2153(2001).
 RL FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of the phosphopantetheine prosthetic group from ACP (By similarity).
 CC -1- CATALYTIC ACTIVITY: Holo-[acyl-carrier protein] + H(2)O = 4'-phosphopantetheine + apo-[acyl-carrier protein].
 CC -1- SIMILARITY: Belongs to the acpd family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL445564; CAC13485.1; -.
 DR PIR; H90550; H90550.
 DR MYPULIST; MYPV_3120; -.
 DR HAMAP; MF_01216; -. 1.
 DR InterPro; IPR003680; NADHdh_2.
 DR Pfam; PF02525; Flavodoxin_2; 1.
 DR Hypothetical protein; Hydrolase; Complete proteome.
 KW SEQUENCE 198 AA; 22673 MW; 453B99E001609F04 CRC64;

Query Match 87.5%; Score 42; DB 1; Length 198;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWYPM 6
 |||||
 DB 139 GWYPM 143

RESULT 4
 O69A.DROME STANDARD; PRT; 393 AA.
 ID O69A.DROME
 AC P82985;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative odorant receptor 69a.
 GN OR69A.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.C., Mortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Baerentzen R.M., Basu P., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.A., Bouck J., Brockstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclio J.M.,
 RA Palazolo R., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Palzer K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT The genome sequence of *Drosophila melanogaster*.
 RL Science 287:2185-2195(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Robertson H.M.;
 RL Unpublished observations (May-2001).
 CC -1- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT
 CC RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AEO03539; -; NOT ANNOTATED_CDS.
 DR FlyBase: FBgn0041622; Or69A.
 DR InterPro: IPR004117; 7tm_6.
 DR Pfam: PF02949; 7tm_6; 1.
 KW Hypothetical protein; Transmembrane; G-protein coupled receptor;
 KW Glycoprotein; Olfaction; Multigene family.
 FT DOMAIN 1 39
 FT TRANSMEM 40 60
 FT DOMAIN 61 69
 FT TRANSMEM 70 90
 FT DOMAIN 91 138
 FT TRANSMEM 139 159
 FT DOMAIN 160 208
 FT TRANSMEM 209 229
 FT DOMAIN 230 269
 FT TRANSMEM 270 290
 FT DOMAIN 291 305
 FT TRANSMEM 306 326
 FT DOMAIN 327 365
 FT TRANSMEM 366 386

FT DOMAIN 387 393 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 393 AA; 46096 MW; EC1EFB9115362C CRC64;
 Query Match 83.3%; Score 40; DB 1; Length 393;
 Best local similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 NGWPW 6
 Db 176 NTWPW 181
 RESULT 5
 YAO_C_SCHPO STANDARD; PRT; 371 AA.
 ID YAO_C_SCHPO
 AC 009680;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein CSH10.12c in chromosome I.
 GN SPKCSH10.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID:4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE-91848401; PubMed-11859360;
 RA Wood V., Gwilliam R., Rayandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Brooks K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeert G., Aert R., Robben J., Grymprez B.,
 RA Wellens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gebel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of *Schizosaccharomyces pombe*."
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: TO S.POMBE SPBC4C3.08 AND SPBC4C3.09.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z49811; CA89962.1; -;
 DR PIR: T36976; S55490.
 DR GeneDB-Spombe; SPKCSH10.12c; -;
 DR InterPro: IPR002495; Glyco_transf_8.
 DR Pfam: PF01501; Glyco_transf_8; 1.

KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 17 33 POTENTIAL.
 SQ SEQUENCE 371 AA; 43688 MW; 890BADADA0E66379B CRC64;

Query Match 81.2%; Score 39; DB 1; Length 371;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 DB 307 GDMFPM 312

RESULT 6

PME_DAUCA STANDARD; PRT; 319 AA.
 AC P83218;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Pectinesterase (EC 3.1.1.11) (Pectin methylsterase) (PE).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE.

RC STRAIN=cv. Tiptop; TISSUE=Root;
 RX MEDLINE=21960069; PubMed=11964128;
 RA Martovic O., Cederlund E., Griffiths W.J., Lipka T., Joernvall H.;
 RT "Characterization of carrot pectin methylsterase";
 RL Cell. Mol. Life Sci. 59:513-518(2002).
 CC -1- FUNCTION: Catalyzes the deesterification of methyl-esterified D-
 galactosiduronic acid units in pectic compounds. It participates
 in modulating cell wall during fruit ripening, cell wall extension
 during pollen germination, and in defense mechanisms against
 pathogens.
 CC -1- CATALYTIC ACTIVITY: Pectin + N H(2)O = N methanol + pectate.
 CC -1- SIMILARITY: BELONGS TO THE PECTINESTERASE FAMILY.
 DR PDB: 1Q08; 18-APR-02.

DR InterPro: IPR000070; Pectinesterase.
 DR Pfam: PF01095; Pectinesterase; 1.
 DR PROSITE: PS00800; PECTINESTERASE_1; 1.
 DR PROSITE: PS00503; PECTINESTERASE_2; 1.
 KM Hydrolyse: Asparyl esterase; Cell wall; Pyrrolidone carboxylic acid;
 KM 3d-structure.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT ACT_SITE 136 136 BY SIMILARITY.
 FT ACT_SITE 157 157 BY SIMILARITY.
 SQ SEQUENCE 319 AA; 34254 MW; 359675FF36FD7625 CRC64;

Query Match 79.2%; Score 38; DB 1; Length 319;
 Best Local Similarity 80.0%; Pred. No. 35;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
 DB 248 GWFPM 252

RESULT 7

GUNH_CLOTM STANDARD; PRT; 900 AA.
 ID GUNH_CLOTM
 AC P16218;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Endoglucanase H precursor (EC 3.2.1.4) (EGH) (Endo-1,4-beta-glucanase)
 DE (Cellulase H).
 GN CELH.
 OS Clostridium thermocellum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.
 OX NCBI_TaxID=1515;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIB 10682;
 RX MEDLINE=90323606; PubMed=2197182;
 RA Yagoue E., Beguin P., Aubert J.-P.;
 RT "Nucleotide sequence and deletion analysis of the cellulase-encoding
 gene celH of Clostridium thermocellum.";
 RL Gene 89:61-67(1990)

CC -1- FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-
 GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-
 GLUCANS.

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 linkages in cellulose, lichenin and cereal beta-D-glucans.
 CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
 WELL AS IN OTHER C.THERMOCCELLUM CELLULOSE ENZYMS. THIS DOMAIN
 MAY FUNCTION AS THE BINDING LIGAND FOR THE ST COMPONENT.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY
 A (FAMILY 5 OF GLYCOSYL HYDROLASES).
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 26 OF
 GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC EMBL: M31903; AAA23225.1; -.
 DR PIR: JH0157; JH0157.
 DR HSSP: P07985; ICEC.
 DR InterPro: IPR005087; CBM_11.
 DR InterPro: IPR002105; Dockerin_1.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001547; Glyco_hydro_5.
 DR Pfam: PF03425; CBM_11; 1.
 DR Pfam: PF00150; cellulase; 1.
 DR Pfam: PF00404; Dockerin_1; 2.
 DR PROSITE: PS00018; EF HAND; UNKNOWN; 1.
 DR PROSITE: PS00448; CLOS_CELLULOSE_RPT; 2.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_5; 1.
 KM Cellulose degradation; Hydrolyase; Glycosidase; Repeat; Signal.
 FT SIGNAL 1 44
 FT CHAIN 45 900
 FT DOMAIN 45 630
 FT DOMAIN 631 654
 FT DOMAIN 655 900
 FT ACT_SITE 460 460
 FT ACT_SITE 565 565
 FT ACT_SITE 833 895
 FT DOMAIN 833 895
 FT REPEAT 833 856
 FT REPEAT 872 895
 SQ SEQUENCE 900 AA; 102415 MW; 973AEP1954FC246B CRC64;

Query Match 78.1%; Score 37.5; DB 1; Length 900;
 Best Local Similarity 85.7%; Pred. No. 16+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 NG-WYPM 6
 DB 133 NGDWYPM 139

RESULT 8

BPT1_YEAST STANDARD; PRT; 1559 AA.
 ID BPT1_YEAST
 AC P14772;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Bile pigment transporter 1.
GN BPT1 OR YLL015W OR L1313.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / FY23;
RX MEDLINE=66405918; PubMed=8810043;
RA Miosga T., Zimmermann F.K.;
RT "Sequence analysis of the GEN2 region of Saccharomyces cerevisiae on
RT a 4.3 kb fragment of chromosome XII including an open reading frame
RT homologous to the human cystic fibrosis transmembrane conductance
RT regulator protein CFTR.";
RL Yeast 12:693-708(1996).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=5288c;
RX MEDLINE=97197984; PubMed=9046100;
RA Purnelle B., Goffeau A.;
RT "The sequence of 32kb on the left arm of yeast chromosome XII reveals
RT six known genes, a new member of the serpinuperin family and a new
RT ABC transporter homologous to the human multidrug resistance
RT protein.";
RL Yeast 13:183-188(1997).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansgorge W.,
RA Jones V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Britlan K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Nittel D., Hilbert H., Hilger F., Kleine K., Koettler P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Westl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetle D., Purnelle B., Rechmann S., Rieger M., Rink M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verlaesselt P.,
RA Vlerendeels F., Voet M., Volckaert G., Voss H., Wambuit R., Wedler E.,
RA Medler H., Zimmermann F.K., Zollner A., Hant J., Holsel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN (4)
RP SEQUENCE OF 1-245 FROM N.A.
RX MEDLINE=89306677; PubMed=2545538;
RA Boy-Marcotte E., Damak F., Camonis J., Garreau H., Jaquet M.;
RT "The C-terminal part of a gene partially homologous to CDC 25 gene
RT suppresses the cdc25-5 mutation in Saccharomyces cerevisiae.";
RL Gene 77:21-30(1989).
RN (5)
RP FUNCTION.
RX MEDLINE=20253522; PubMed=10790694;
RA Petrovic S., Pascolo L., Gallo R., Cupelli F., Ostrow J.D.,
RA Goffeau A., Tirlbelli C., Bruschl C.V.;
RT "The products of YCF1 and YLL015W (BPT1) cooperate for the ATP-
RT dependent vacuolar transport of unconjugated bilirubin in
RT Saccharomyces cerevisiae.";
RL Yeast 16:561-571(2000).
RN (6)
RP FUNCTION: Cooperates for the ATP-dependent vacuolar transport of
RN bilirubin.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X97560; CAA66162.1; -

DR EMBL: Z73120; CAA97460.1; -;
DR EMBL: X91488; CAA62776.1; -;
DR EMBL: M26647; AAA16564.1; -;
DR PIR: S64757; S64757.
DR HSSP: P13569; INBD.
DR SGD: S0003938; BPT1.
DR GO: GO:0000329; C:Vacuolar membrane (sensu Fungi); IDA.
DR GO: GO:0015127; F:Bilirubin transporter activity; IGI.
DR GO: GO:0015086; F:cadmium ion transporter activity; IDA.
DR GO: GO:0015723; P:bilirubin transport; IGI.
DR GO: GO:0015691; P:cadmium ion transport; IDA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Transmembrane; Glycoprotein; Transport.
FT DOMAIN 1 29
FT TRANSMEM 30 50
FT DOMAIN 51 84
FT TRANSMEM 85 105
FT DOMAIN 106 110
FT TRANSMEM 111 127
FT DOMAIN 128 139
FT TRANSMEM 140 160
FT DOMAIN 161 178
FT TRANSMEM 179 199
FT DOMAIN 200 283
FT TRANSMEM 284 304
FT DOMAIN 305 333
FT TRANSMEM 334 354
FT DOMAIN 355 410
FT TRANSMEM 411 431
FT DOMAIN 432 434
FT TRANSMEM 435 455
FT DOMAIN 456 518
FT TRANSMEM 519 539
FT DOMAIN 540 560
FT TRANSMEM 561 581
FT DOMAIN 582 592
FT TRANSMEM 593 993
FT DOMAIN 994 1030
FT TRANSMEM 1031 1052
FT DOMAIN 1053 1095
FT TRANSMEM 1096 1116
FT DOMAIN 1117 1117
FT TRANSMEM 1118 1138
FT DOMAIN 1139 1209
FT TRANSMEM 1210 1230
FT DOMAIN 1231 1235
FT TRANSMEM 1236 1256
FT DOMAIN 1257 1559
FT NP_BIND 672 679
FT NP_BIND 1336 1343
FT CARBOHYD 1011 1011
SQ SEQUENCE 1559 AA; 176873 MW; 0460F561E3125D5 CRC64;
N-LINKED (GLCNAC...) (POTENTIAL).
Query Match 77.1%; Score 37; DB 1; Length 1559;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NGWP 5
Db 213 NGWYP 217
RESULT 9
Y017_METUA STANDARD; PRT; 214 AA.

```

AC 060323;
DE 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ0017.
GN MJ0017.
OS Methanococcus jannaschii.
OC Archaeae; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierkegaard A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0139.1 AND MJ1466.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U67460; AAB97994.1; -
CC PIR: A64302; A64302.
CC TIGR: MJ0017; -
CC InterPro: IPR001584; Rve.
CC Pfam: PF00665; rve; 1.
CC Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 214 AA; 25506 MW; 0F135EC748F933D CRC64;

Query Match 75.0%; Score 36; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
DB 146 WYPM 149

RESULT 10
YF66_METJA STANDARD: PRT; 214 AA.
ID YF66_METJA
AC Q58861;
DE 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ1466.
GN MJ1466.
OS Methanococcus jannaschii.
OC Archaeae; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierkegaard A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0139.1 AND MJ1466.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U67460; AAB97994.1; -
CC PIR: A64302; A64302.
CC TIGR: MJ0017; -
CC InterPro: IPR001584; Rve.
CC Pfam: PF00665; rve; 1.
CC Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 214 AA; 25506 MW; 0F135EC748F933D CRC64;

```

```

RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0139.1 AND MJ1466.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U67587; AAB99474.1; -
CC PIR: A64483; A64483.
CC TIGR: MJ1466; -
CC InterPro: IPR001584; Rve.
CC Pfam: PF00665; rve; 1.
CC Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 214 AA; 25626 MW; 3C6D5A76051F17EB CRC64;

Query Match 75.0%; Score 36; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
DB 146 WYPM 149

RESULT 11
R18B_MOUSE STANDARD: PRT; 254 AA.
ID R18B_MOUSE
AC Q99N84; Q9CRK0; Q9DCR8;
DE 28-FEB-2003 (Rel. 41, Created)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE 28S ribosomal protein S18b, mitochondrial precursor (MRP-S18-b)
DE (MRP-S18b) (MRP-S18-2).
GN MRP-S18b.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC MEDLINE=21413863; PubMed=11402041;
RX Suzuki T., Terasaki M., Takemoto-Hori C., Hanada T., Ueda T., Wada A.,
RA Watanabe K.;
RT "Proteomic analysis of the mammalian mitochondrial ribosome.
RT Identification of protein components in the 28 S small subunit."
RL J. Biol. Chem. 276:33181-33195(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells, and kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.B., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,

```


RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Togo-Oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Collins E.A., Grouse L.H., Derge J.G.,
RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmechel A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named Isoforms=2;
CC Name=1;
CC IsoId=Q09N84-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q09N84-2; Sequence=VSP_005723;
CC Note-No experimental confirmation available;
CC -1- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 19.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB049954; BAB41007.1; -;
DR EMBL; AK002554; -; NOT_ANNOTATED_CDS.
DR EMBL; AK010250; -; NOT_ANNOTATED_CDS.
DR EMBL; BC021752; AAH21752.1; -;
DR MGD; MGI:1914223; Mrp818b.
DR GO; GO:0005763; C:mitochondrial small ribosomal subunit; ISS.
DR GO; GO:0003735; F:structural constituent of ribosome; ISS.
DR GO; GO:0006412; P:protein biosynthesis; ISS.
DR InterPro; IPR001648; Ribosomal_S18.
DR Pfam; PF01084; Ribosomal_S18; 1.
DR PROSITE; PS00057; RIBOSOMAL_S18; FALSE_NEG.
KW Ribosomal protein; Mitochondrion; Transit peptide;
KW Alternative splicing.
FT TRANSIT 1
FT CHAIN ? 254 MITOCHONDRION (BY SIMILARITY).
FT VASPLIC 25 116 Missing (in isoform 2).
FT FT
FT VASPLIC 25 116 /Ftrid-VSP_005723
SQ SEQUENCE 254 AA; 28702 MW; 2284983FD5848791 CRC64;
Query Match 75.0%; Score 36; DB 1; Length 254;

Best Local Similarity 100.0%; Pred. No. 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 WYFW 6
Db 197 WYFW 200
RESULT 12
RIBB_HUMAN STANDARD; PRT; 258 AA.
ID RIBB_HUMAN
AC Q9Y676; Q9BS27;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 28S ribosomal protein S18b, mitochondrial precursor (MRP-S18-B)
DE (Mrp18b) (MRP-S18-2) (PTD017 protein) (HSPC183).
GN MRP518B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Plutitary tumor;
RA Zhang Q.-H., Guan Z.-Q., Dai M., Song H., Mao Y.F., Wu X.Y., Mao M.,
RA Fu G., Luo M., Chen J.H., Hu R.,
RT "Human PRD017 gene",
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.,
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.",
RL Genome Res. 10:1546-1560(2000).
RN [3]
RP SEQUENCE FROM N.A.
RP TISSUE=Heart;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmechel A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP IDENTIFICATION.
RX MEDLINE=21276436; PubMed=11279123;
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.,
RT "The small subunit of the mammalian mitochondrial ribosome:
RT identification of the full complement of ribosomal proteins present.",
RL J. Biol. Chem. 276:19363-19374(2001).
CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.

```

CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF100761; AAD43025.1; -
DR EMBL: AF151017; AAF36103.1; -
DR EMBL: BC005373; AAH05373.1; -
DR Genew: HGNC:14516; MRPS18B.
DR GO: GO:0005763; C:mitochondrial small ribosomal subunit; NAS.
DR GO: GO:0003735; F:structural constituent of ribosome; NAS.
DR GO: GO:0006412; P:protein biosynthesis; NAS.
DR InterPro: IPR010648; Ribosomal_S18.
DR Pfam: PF01084; Ribosomal_S18; 1.
DR PROSITE: PS00057; RIBOSOMAL_S18; FALSE NEG.
DR Ribosomal protein; Mitochondrion; Transit peptide.
FT TRANSIT 1 258 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 196 196 28S RIBOSOMAL PROTEIN S18B.
FT CONFLICT 196 196 G -> S (IN REF. 3).
SQ SEQUENCE 258 AA; 29395 MW; B4C83E5593796C5D CRC64;

Query Match 75.0%; Score 36; DB 1; Length 258;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
DB 199 WYPM 202

RESULT 13
CP21_HORVU
ID CP21_HORVU STANDARD; PRT; 324 AA.
AC P55747;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-1 precursor (EC 3.4.16.6) (CP-MIT.1)
DE (Fragment).
GN CXP;2-1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_Taxid=4513;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Alexis; TISSUE=grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Degon F., Rocher A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
RT germination of the barley grain."
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
CC -1- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
CC or lysine residue.
CC -1- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
CC COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW
CC LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE
CC ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -1- PTM: THE LINKER PEPTIDE IS ENDOPEPTIDOLYTIALLY EXCISED DURING
CC ENZYME MATURATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X78876; CAB58992.1; -
DR HSSP: P08819; 1WHT.
DR MEROPS: S10.005; -.
DR InterPro: IPR001563; Serine_carpept.
DR Pfam: PF00450; serine_carpept. 1.
DR PRINTS: PR00724; CRBOXYPRASEC.
DR ProDom: PD001189; Serine_carpept. 1.
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER. 1.
DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS. 1.
DR Hydroxylase; Carboxypeptidase; Glycoprotein; Zymogen; Multigene family.
FT NON_TER 1 1
FT CHAIN 1 149 SERINE CARBOXYPEPTIDASE II-1, CHAIN A.
FT PROPEP 150 162 LINKER PEPTIDE (BY SIMILARITY).
FT CHAIN 163 324 SERINE CARBOXYPEPTIDASE II-1, CHAIN B.
FT ACT_SITE 41 41 BY SIMILARITY.
FT ACT_SITE 239 239 BY SIMILARITY.
FT ACT_SITE 291 291 BY SIMILARITY.
FT DISULFID 109 121 BY SIMILARITY.
FT DISULFID 145 170 BY SIMILARITY.
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 324 AA; 37408 MW; 70F081D6B9723A60 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 324;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WYPM 6
DB 262 WYPM 265

RESULT 14
HOLB_BUCAP
ID HOLB_BUCAP STANDARD; PRT; 328 AA.
AC O8K012;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase III, delta' subunit (EC 2.7.7.7).
DE HOLB OR BUSG342.
GN Buchnera aphidicola (subsp. Schizaphis graminum).
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_Taxid=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tams J., Klasson L., Canback B., Neeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria."
RL Science 296:2376-2379(2002).
CC -1- FUNCTION: DNA polymerase III is a complex, multichain enzyme
CC responsible for most of the replicative synthesis in bacteria.
CC This DNA polymerase also exhibits 3' to 5' exonuclease activity
CC (by similarity).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + [DNA](n).
CC -1- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIII' complex. POLIII' associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE014110; AAM67896.1; .
 KM Transferase; DNA-directed DNA polymerase; DNA replication;
 KW Complete proteome.
 SQ SEQUENCE 328 AA; 38927 MW; 160BA0FE21BAD08D CRC64;
 Query Match 75.0%; Score 36; DB 1; Length 328;
 Beet Local Similarity 100.0%; Pred. No. 72;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 WYPM 6
 DB 3 WYPM 6
 RESULT 15
 ID HOLB_ECOLI STANDARD; PRT; 334 AA.
 AC P28631.
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA polymerase III, delta' subunit (EC 2.7.7.7).
 GN HOLB OR B1099.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA MEDLINE-93280136; PubMed-8505303;
 RA Dong Z., Onrust R., Skangalis M., O'Donnell M.;
 RT "DNA polymerase III accessory proteins. I. holA and holB encoding
 RT delta and delta'.";
 RL J. Biol. Chem. 268:11758-11765(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAF102;
 RA Carter J.R., Franden M.A., Hebersold R.H., McHenry C.S.;
 RT "Identification, isolation, and characterization of the structural
 RT gene encoding the delta' subunit of Escherichia coli DNA polymerase
 RT III holoenzyme.";
 RL J. Bacteriol. 175:3812-3822(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA MEDLINE-97061202; PubMed-8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 RN [5]

RP CHARACTERIZATION.
 RX MEDLINE-93280137; PubMed-8505304;
 RA Onrust R., O'Donnell M.;
 RT "DNA polymerase III accessory proteins. II. Characterization of delta
 RT and delta'.";
 RL J. Biol. Chem. 268:11766-11772(1993).
 RN [6]
 RP REVIEW.
 RX MEDLINE-92246902; PubMed-1575709;
 RA O'Donnell M.;
 RT "Accessory protein function in the DNA polymerase III holoenzyme from
 RT E. coli.";
 RL Bioessays 14:105-111(1992).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE-98028572; PubMed-9363942;
 RA Guenther B., Onrust R., Sall A., O'Donnell M., Kurlyan J.;
 RT "Crystal structure of the delta' subunit of the clamp-loader complex
 RT of E. coli DNA polymerase III.";
 RL Cell 91:335-345(1997).
 CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
 CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
 CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
 CC + (DNA)(N).
 CC -1- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
 CC epsilon and theta chains) that associates with a tau subunit. This
 CC core dimerizes to form the PolIII' complex. PolIII' associates
 CC with the gamma complex (composed of gamma, delta, delta', psi and
 CC chi chains) and with the beta chain to form the complete DNA
 CC polymerase III complex. The final composition of the complex is:
 CC (alpha, epsilon, theta)[2]-tau[2]-(gamma, delta, delta', psi, chi)[2]-
 CC beta[4].
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L04577; AAA23708.1; .
 DR EMBL: L01483; AAA23696.1; .
 DR EMBL: AE000210; AAC74183.1; .
 DR EMBL: D90745; BAA35906.1; .
 DR PIR: S35523; S35523.
 DR PDB: 1A5T; 27-MAY-98.
 DR Ecogene; Egl1500; holB.
 DR InterPro: IPR004622; HolB.
 DR InterPro: IPR000862; Rctdomain.
 DR TIGRFAMS: TIGR00678; holB; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW 3D-structure; Complete proteome.
 FT CONFLICT 166 166 P -> G (IN REF. 1).
 FT HELIX 5 7
 FT HELIX 8 19
 FT TURN 20 21
 FT STRAND 26 30
 FT TURN 32 34
 FT HELIX 37 48
 FT TURN 49 49
 FT STRAND 54 54
 FT STRAND 55 56
 FT STRAND 57 57
 FT HELIX 63 70
 FT TURN 71 71
 FT TURN 74 75
 FT STRAND 76 79
 FT TURN 83 84
 FT STRAND 88 88
 FT HELIX 90 99
 FT TURN 100 101

FT TURN 105 106
FT STRAND 110 114
FT HELIX 117 119
FT STRAND 120 120
FT STRAND 122 132
FT HELIX 133 133
FT TURN 137 138
FT STRAND 139 145
FT HELIX 148 150
FT HELIX 153 156
FT TURN 157 158
FT STRAND 160 163
FT HELIX 169 179
FT HELIX 184 193
FT TURN 194 196
FT HELIX 198 203
FT TURN 204 205
FT HELIX 208 226
FT HELIX 230 232
FT HELIX 233 236
FT TURN 239 240
FT HELIX 241 255
FT TURN 256 269
FT TURN 268 280
FT HELIX 271 280
FT HELIX 283 303
FT HELIX 308 322
FT TURN 323 323
FT TURN 325 326
SQ SEQUENCE 334 AA; 36936 MW; DC9DA644AD8A096 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WYPW 6
|||
Db 3 WYPW 6

Search completed: August 20, 2003, 12:34:59
Job time : 3.38822 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 11.3494 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-32
Perfect score: 48
Sequence: 1 NGWYPW 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	42	87.5	212	16	08EWL4	08EWL4 mycoplasma
2	42	87.5	3436	5	08SVL2	08SVL2 encephalito
3	40	83.3	345	12	05S837	05S837 yaounde vir
4	40	83.3	414	5	0810J1	0810J1 dirosophila
5	40	83.3	456	2	09KK86	09KK86 brevbacter
6	40	83.3	919	15	09WPP0	09WPP0 chimpanzee
7	39	81.2	708	5	093744	093744 caenorhabdi
8	39	81.2	767	17	0973G2	0973G2 sulfolobus
9	38	79.2	99	10	08H223	08H223 populus x c
10	38	79.2	158	16	08BKV7	08BKV7 xanthomonas
11	38	79.2	171	16	08P8T2	08P8T2 xanthomonas
12	38	79.2	223	10	09LPX8	09LPX8 arabidopsis
13	38	79.2	345	16	0987E1	0987E1 rhizobium l
14	38	79.2	376	3	043062	043062 schizosacch
15	38	79.2	474	17	08TX05	08TX05 methanopyru
16	38	79.2	557	10	094C39	094C39 arabidopsis

17	38	79.2	764	17	097Z77	097Z77 sulfolobus
18	37.5	78.1	540	16	09HWI3	09HWI3 pseudomonas
19	37.5	78.1	542	10	08LRT4	08LRT4 arabidopsis
20	37.5	78.1	544	10	09LOS1	09LOS1 arabidopsis
21	37.5	78.1	552	10	09FEE6	09FEE6 arabidopsis
22	37	77.1	145	16	08KCB9	08KCB9 chlorobium
23	37	77.1	206	10	0946V3	0946V3 zea mays (m
24	37	77.1	232	2	093543	093543 neisseria m
25	37	77.1	235	2	093545	093545 neisseria m
26	37	77.1	238	2	093544	093544 neisseria m
27	37	77.1	248	2	093542	093542 neisseria m
28	37	77.1	250	2	051228	051228 neisseria m
29	37	77.1	252	2	09AE79	09AE79 neisseria m
30	37	77.1	262	16	09X8R8	09X8R8 streptomyce
31	37	77.1	272	2	051230	051230 neisseria m
32	37	77.1	272	16	051229	051229 neisseria m
33	37	77.1	272	16	051227	051227 neisseria m
34	37	77.1	287	16	08Y2A8	08Y2A8 ralstonia s
35	37	77.1	314	16	08F8F5	08F8F5 escherichia
36	37	77.1	349	11	08BM15	08BM15 mus musculu
37	37	77.1	352	16	P73267	P73267 synchocyst
38	37	77.1	414	16	08YXJ8	08YXJ8 ralstonia s
39	37	77.1	606	16	0988W4	0988W4 rhizobium l
40	37	77.1	768	5	09VGL9	09VGL9 dirosophila
41	37	77.1	1230	16	08FIF6	08FIF6 escherichia
42	37	77.1	1236	2	0939T3	0939T3 bacillus th
43	37	77.1	1690	5	09YXU0	09YXU0 dirosophila
44	37	77.1	1708	5	016039	016039 dirosophila
45	37	77.1	1875	5	093691	093691 caenorhabdi

ALIGNMENTS

RESULT 1
ID 08EWL4 PRELIMINARY; PRT; 212 AA.
AC 08EWL4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Acyl carrier protein phosphodiesterase.
GN MYPEL890.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=1246555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.,
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL: AP004170; BAC43980.1; -;
KW Complete proteome.
SQ SEQUENCE 212 AA; 23955 MW; 5C54F59C3C816PDF CRC64;
Query Match 87.5%; Score 42; DB 16; Length 212;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
ID 08SVL2 PRELIMINARY; PRT; 3436 AA.
AC 08SVL2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein ECU05_0540.
 GN ECU05_0540.
 OS Eucephalitozoon cuniculi.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Eucephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA MEDLINE=21576510; PubMed=11719806;
 RA Katsuka M.D., Duprat S., Cornillot E., Metenier G., Thomarar F.,
 RA Prestier G., Barbe V., Peyretallade E., Brothier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gony M.,
 RA Weissenbach J., Vives C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Eucephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 DR EMBL; AL590445; CAD26573.1; -.
 DR InterPro; IPR003151; FAT.
 DR Pfam; PF02259; FAT; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 3436 AA; 395116 MW; E5F687024463851A CRC64;
 OY 2 GWYPM 6
 DB 2140 GWYPM 2144
 RESULT 3
 ID 055837 PRELIMINARY; PRT; 345 AA.
 AC 055837.
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NS5 protein (Fragment).
 GN NS5.
 OS Yaounde virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=64319;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DAKAY 276;
 RX MEDLINE=98080391; PubMed=9420202;
 RX Kuno G., Chang G.J., Tsuchiya K.R., Karabatsos N., Cropp C.B.;
 RT "Phylogeny of the genus Flavivirus.";
 RL J. Virol. 72:73-83(1998).
 DR EMBL; AF013413; AAC58801.1; -.
 DR InterPro; IPR00208; Flavivirus.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVIT.
 DR Pfam; PF00972; Flavivirus; 1.
 DR PROSITE; PS50507; RDRP_POSITIVE; 1.
 DR PROSITE; PS50521; RDRP_VIRAL; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 345 AA; 39245 MW; DC667CC97761AB5D CRC64;
 Query Match 83.3%; Score 40; DB 12; Length 345;
 Best Local Similarity 83.3%; Pred. No. 1.le+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 DB 241 NGWYPM 246
 RESULT 4
 ID 0810J1 PRELIMINARY; PRT; 414 AA.
 AC 0810J1.
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE CG32116-PB.
 GN CG32116.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OX Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Franko C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Boucek J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler A., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Talali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson J.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodgerf, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Bankon J., An H., Baldwin D., Bankon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Drenek D., Farfan D.,
 RA Fertiera S., Frise E., Galle R.F., Gary N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Paclebo J., Parags V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celisner S.E.,
 RA Camp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richner J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RP EMBL: AE003539; AAN1843.1; -
 SO SEQUENCE 414 AA; 48538 MW; 5488B0A191184D06 CRC64;

Query Match 83.3%; Score 40; DB 5; Length 414;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 DB 176 NTWYPM 181

RESULT 5
 O9KK86 PRELIMINARY; PRT; 456 AA.
 AC O9KK86;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Hypothetical 50.2 kDa protein (Fragment).
 OS Brevibacterium linens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Brevibacteriaceae; Brevibacterium.
 OX NCBI_TaxID=1703;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 20426;
 RX MEDLINE=20279196; PubMed=10821176;
 RA Krubasik P., Sandmann G.;
 RT "A carotenogenic gene cluster from *Brevibacterium linens* with novel
 RT lycopene cyclase genes involved in the synthesis of aromatic
 RT carotenoids";
 RL Mol. Gen. Genet. 263:423-432(2000).
 DR EMBL: AF139916; AAF65580.1; -
 KW Hypothetical protein.
 FT NON TER
 SO SEQUENCE 456 AA; 50156 MW; E35B39F2D7CA4CB8 CRC64;

Query Match 83.3%; Score 40; DB 2; Length 456;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 DB 333 NGWYPM 338

RESULT 6
 O9WPP0 PRELIMINARY; PRT; 919 AA.
 ID O9WPP0

AC O9WPP0;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SVSUN;
 RX MEDLINE=99370217; PubMed=10438863;
 RA Beer B.E., Baller E., Goeken R., Dapolito G., Coulibaly C., Norley S.,
 RA Kurth R., Gautier J.P., Gautier-Hon A., Vallier D., Sharp P.M.,
 RA Hirsch V.M.;
 RT "Simian immunodeficiency virus (SIV) from sun-tailed monkeys
 RT (Cercopithecus solatus): evidence for host-dependent evolution of SIV
 RT within the C. lhoesti superspecies";
 RL J. Virol. 73:7734-7744(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SVSUN;
 RA Beer B.E., Hirsch V.M.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF131870; AAD39758.1; -
 DR InterPro: IPR000328; Env.GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
 SO SEQUENCE 919 AA; 104803 MW; 57682041B7FDEDE9 CRC64;

Query Match 83.3%; Score 40; DB 15; Length 919;
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 DB 444 NKWYPM 449

RESULT 7
 O93744 PRELIMINARY; PRT; 708 AA.
 AC O93744;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE F48C11.2 protein.
 GN F48C11.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wild A.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z80789; CAB02551.1; -
 DR WormPep: F48C11.2; CE10750.
 DR InterPro: IPR001304; LECTIN_C.
 DR Pfam: PF00059; LECTIN_C; 1.
 DR SMART: SM00034; CLECT_1.
 SO SEQUENCE 708 AA; 75157 MW; D7EF7ED81B1A1A9 CRC64;

Query Match 81.2%; Score 39; DB 5; Length 708;
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 |||
 Db 696 NGMKPW 701

RESULT 8

OY973G2 PRELIMINARY; PRT; 767 AA.
 AC O973G2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative oligosaccharyl transferase.
 GN ST0940.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC NCBL_TaxID=111955;
 OX NCBL_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed-11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa J., Takamida M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT *Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000984; BAB5951.1; -;
 DR InterPro: IPR001092; HLH_basic.
 DR PROSITE: PS00038; HLH_1; 1.
 KW Transferrase; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 767 AA; 83742 MW; 277493A13462EC18 CRC64;

Query Match 81.2%; Score 39; DB 17; Length 767;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 :||:|
 Db 74 HGMPW 79

RESULT 9

OY8H223 PRELIMINARY; PRT; 99 AA.
 AC O8H223;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative pectin methyltransferase (Fragment).
 GN PME.
 OS Populus x canadensis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Malpighiales; Salicaceae; Populus.
 OX NCBL_TaxID=80863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lapante P., Seguin A.;
 RT "Differentially expressed genes following elicitor treatments in
 RT poplar";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY129243; AAN05419.1; -;
 FT NON_TER 1
 SQ SEQUENCE 99 AA; 10555 MW; 93E44F3B89E506BA CRC64;

Query Match 79.2%; Score 38; DB 10; Length 99;
 Best Local Similarity 80.0%; Pred. No. 68;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYPM 6
 ||:|
 Db 25 GWMPW 29

RESULT 10

OY8PKV7 PRELIMINARY; PRT; 158 AA.
 AC O8PKV7;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein XAC2050.
 GN XAC2050.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBL_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-306 / ATCC 13902 / XV 101;
 RX MEDLINE-22022145; PubMed-12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL: AE011840; AAM36912.1; -;
 DR InterPro: IPR001412; tRNA-synt_1.
 DR PROSITE: PS00178; AA_tRNA_LIGASE_I; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 158 AA; 17289 MW; C1B1FB23A6E930B CRC64;

Query Match 79.2%; Score 38; DB 16; Length 158;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NGWYPM 6
 :||:|
 Db 138 HGMPW 143

RESULT 11

OY8P8T2 PRELIMINARY; PRT; 171 AA.
 AC O8P8T2;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein XCC2157.
 GN XCC2157.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBL_TaxID=340;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33913 / NCPPB 528;
 RX MEDLINE-22022145; PubMed-12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio L.M.C., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
 RA Spilola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezsa R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitchjama J.P.,
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 DR EMBL: AE012322; AAM41437.1; -
 DR InterPro: IPR001412; TRNA-synL.I.
 DR PROSITE: PS00178; AA:TRNA_LIGASE_I.1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 171 AA; 18661 MW; B9C10E1FA22051D CRC64;

Query Match 79.2%; Score 38; DB 16; Length 171;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 151 HGWPM 156

RESULT 12
 Q9LPX8 PRELIMINARY; PRT; 223 AA.
 AC Q9LPX8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE T23J18.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 RA Kim C., Altieri H., Bel O., Chin C., Chou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 RA Leitz C., Li J., Liu A., Liu K., Liu S., Mukharsy N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thavavel A.,
 RA Tortum M., Vaysberg M., Yu G., Federpiel N.A., Theologis A.,
 RA Ecker J.R.,
 RT "Genomic sequence for Arabidopsis thaliana BAC T23J18 from chromosome
 RT 1."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011661; AAF16637.1; -
 DR InterPro: IPR000070; Pectinesterase.
 DR Pfam: PF01095; Pectinesterase; 1.
 DR PROSITE: PS00503; PECTINESTERASE_2; 1.
 SO SEQUENCE 223 AA; 25031 MW; 289B968B77307882 CRC64;

Query Match 79.2%; Score 38; DB 10; Length 223;
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GWYPM 6

DB 152 GWPM 156

RESULT 13
 Q987E1 PRELIMINARY; PRT; 345 AA.
 ID Q987E1;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein ml17088.
 GN ML17088.
 OS Rhizobium loti (Mesorhizobium loti).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Phyllobacteriaceae; Mesorhizobium.
 CX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAFF30309;
 RX MEDLINE-21062930; PubMed-11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003011; BAB5362.1; -
 DR InterPro: IPR001173; Glyco.transf.2.
 DR Pfam: PF00535; Glycos.transf.2; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 345 AA; 39450 MW; 131B52D5E56507E1 CRC64;

Query Match 79.2%; Score 38; DB 16; Length 345;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGWYPM 6
 DB 139 NGWYPM 144

RESULT 14
 Q43062 PRELIMINARY; PRT; 376 AA.
 ID Q43062;
 AC Q43062;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE SPBC4C3.09 protein.
 GN SPBC4C3.09.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wood V., Rajadream M.A., Barrell B.G., Lauber J., Hilbert H.,
 RA Duesterhoeft A.,
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL021730; CAA16831.1; -
 DR GeneDB: SPombe; SPBC4C3.09; -
 DR InterPro: IPR002495; Glyco.transf.8.
 DR Pfam: PF01501; Glyco.transf.8; 1.
 SO SEQUENCE 376 AA; 44095 MW; 09D82A6BF65F2400 CRC64;

Query Match 79.2%; Score 38; DB 3; Length 376;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYFW 6
 ||:||
 Db 313 GWFPW 317

RESULT 15

08TX05 PRELIMINARY: PRT: 474 AA.
 AC 08TX05: 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Predicted transposase.
 GN MK0605.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AV19 / DSM 6324 / JCM 9639;
 MEDLINE-21927647; PubMed-11930014;
 RA Stesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Mal'kh A.G., Koonin E.V., Kozhavkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 and monophyly of archaeal methanogens.";
 RL Proc. Natl Acad. Sci. U.S.A. 99:4644-4649 (2002).
 DR EMBL; AE010353; JAM01820.1; -.
 KW Complete proteome.
 SQ SEQUENCE 474 AA; 55032 MW; C24446DBB05B5A91 CRC64;

Query Match 79.2%; Score 38; DB 17; Length 474;
 Best Local Similarity 80.0%; Pred. No. 3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWYFW 6
 ||:||
 Db 179 GWFPW 183

Search completed: August 20, 2003, 12:40:32
 Job time : 13.4605 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 ; Search time 13.1928 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41
Sequence: 1 GCMIPY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	51	23	ABG90939 Human colon specific
2	37	90.2	51	23	ABG04311 Human colon specific
3	36	87.8	49	22	ABG51866 Human liver peptid
4	36	87.8	49	22	ABG31790 Peptide #4441 enco
5	36	87.8	49	22	ABG37019 Peptide #4525 enco
6	36	87.8	49	22	ABG22333 Protein #4332 enco
7	36	87.8	49	22	AA57746 Human brain expres
8	36	87.8	49	22	AAW70159 Human bone marrow
9	36	87.8	49	22	AAW17988 Peptide #4422 enco

10	36	87.8	49	22	AAW30495 Peptide #4532 enco
11	36	87.8	49	22	AAW05626 Peptide #4308 enco
12	36	87.8	49	23	ABG39797 Human peptid enco
13	35	85.4	20	23	AAU71006 M. tuberculosis Rv
14	35	85.4	51	22	AAO11246 Human polypeptide
15	35	85.4	80	24	ABU11869 Human secreted pro
16	35	85.4	109	22	AAO00711 Human polypeptide
17	35	85.4	113	24	ABU71030 Human adipocyte se
18	35	85.4	188	21	AAU71605 Arabidopsis thalia
19	35	85.4	189	21	AAU40279 Arabidopsis thalia
20	35	85.4	191	22	ABG01626 Novel human diagno
21	35	85.4	208	22	AAU55735 Protonbacterium
22	35	85.4	209	22	AAU56988 Beak and feather d
23	35	85.4	220	22	AAU31800 Micromonospora eve
24	35	85.4	248	22	AAU04849 Orthomyxlin biosy
25	35	85.4	248	24	ABP99342 Arabidopsis thalia
26	35	85.4	250	21	AAU71604 Arabidopsis thalia
27	35	85.4	251	21	AAU40278 Arabidopsis thalia
28	35	85.4	260	21	AAU17603 Arabidopsis thalia
29	35	85.4	260	21	AAU40277 Arabidopsis thalia
30	35	85.4	269	21	AAU84911 Amino acid sequenc
31	35	85.4	379	20	AAU37433 Chlamydia trachoma
32	35	85.4	421	20	AAU35514 Chlamydia pneumonia
33	35	85.4	494	21	AAU45012 Maize ZmKCS1 prote
34	35	85.4	505	19	AAU70516 Candida parapsillos
35	35	85.4	505	12	AAU64143 Candida parapsillos
36	35	85.4	507	17	AAU91939 Mouse hippocampus-
37	35	85.4	507	18	AAU26704 Mouse Cyp7b sterol
38	35	85.4	507	22	ABG57286 Mouse ischaemic co
39	35	85.4	800	22	ABG67506 Drosophila melanog
40	35	85.4	974	23	ABP59089 Cell division cycl
41	35	85.4	1330	23	AAU70947 M. tuberculosis Rv
42	34	82.9	23	23	AAU66421 Androgen receptor
43	34	82.9	55	22	AAU42647 Protonbacterium
44	34	82.9	61	22	AAU85213 Human immune/haema
45	34	82.9	68	22	AAU64712 Human immune/haema

ALIGNMENTS

RESULT 1
ABG90939 standard; Protein: 51 AA.
ID ABG90939:
AC
XX 29-NOV-2002 (first entry)
XX
XX Human colon specific polypeptide, #12.
DE
XX Human colon specific polypeptide, #12.
XX
XX Human: colon specific nucleic acid; CSNA: colon specific polypeptide;
KW CSF; vaccine; diagnosis; metastasis; colon cancer; cancer;
KW Immune response; staging; imaging; colorectal cancer; gene therapy;
KW transgenic; engineered colon tissue.
XX
OS Homo sapiens.
XX
XX PN WO200264744-A2.
XX
XX 22-AUG-2002.
XX
XX 12-FEB-2002; 2002MO-US04240.
XX
XX 13-FEB-2001; 2001US-268291P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Piderit A, Macina RA, Sun Y., Liu C;
XX WPI, 2002-657591/70.
XX N-PSDB; ABS67078, ABS67079.
XX

PT New colon specific nucleic acid, useful in gene therapy or as vaccines
 PT for treating colon cancer or colorectal cancer, or non-cancerous colon
 PT diseases, as well as for diagnosing, monitoring or staging these
 PT diseases

XX Claim 11: Page 178; 187pp; English.

CC The invention discloses isolated colon specific nucleic acids (CSNAs) and
 CC the polypeptides (CSPs) that they encode. The colon specific nucleic
 CC acids and polypeptides are useful as vaccines. The colon specific nucleic
 CC acid and polypeptide are also useful for diagnosing and monitoring the
 CC presence and metastases of colon cancer in a patient. The antibody that
 CC specifically binds to the colon specific polypeptide is useful for
 CC determining the presence of a colon specific protein in a sample, as well
 CC as for treating a patient with cancer, particularly by inducing an immune
 CC response against the colon cancer cell expressing the colon specific
 CC nucleic acid molecule or polypeptide. In particular, these colon specific
 CC genes and proteins are useful for identifying, diagnosing, monitoring,
 CC staging, imaging and treating colon cancer (e.g. colorectal cancer) and
 CC non-cancerous disease states in the colon. These are also useful in gene
 CC therapy, production of transgenic animals and cells and in the production
 CC of engineered colon tissue for treatment and research. The sequences
 CC presented in ABG90928-ABG90963 are the human CSPs encoded by the CSNAs.

XX Sequence 51 AA;

Query Match 90.2%; Score 37; DB 23; Length 51;

Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLPY 6
 |||||

DB 43 GGMLPF 48

RESULT 2

ABJ04311

ID ABJ04311 standard; Protein; 51 AA.

XX ABJ04311;

AC ABJ04311;

XX 11-OCT-2002 (first entry)

DE Human colon specific protein SEQ ID NO: 85.

XX Human: colon specific gene; colon specific protein; colon cancer;
 KW colorectal cancer; colon disease; cytostatic; gene therapy.

OS Homo sapiens.

XX WO200242460-A2.

PD 30-MAY-2002.

XX 21-NOV-2001; 2001WO-US43611.

XX 22-NOV-2000; 2000US-252505P.

XX (DIAD-) DIADEXUS INC.

XX Macina RA, Recipon H, Pluta J, Ghosh MG, Sun Y, Liu C;

XX WPI; 2002-583378/62.

PT Novel colon specific polypeptides and polynucleotides useful for
 PT detecting, diagnosing, monitoring, treating, staging and predicting
 PT cancers in humans having cancer and non-cancerous colon disease

PS Claim 11; Page 208; 228pp; English.

CC The present invention provides protein and coding sequences of human
 CC colon specific genes and proteins. These can be used in the treatment of
 CC colonic diseases, including colon and colorectal cancers. The present

CC sequence is a protein of the invention.

XX Sequence 51 AA;

Query Match 90.2%; Score 37; DB 23; Length 51;

Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLPY 6
 |||||

DB 43 GGMLPF 48

RESULT 3

ABG51866

ID ABG51866 standard; Peptide; 49 AA.

XX ABG51866;

XX 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID NO 30514.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO20015273-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver

XX Claim 27; SEQ ID NO 30514; 658pp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 49 AA;

Query Match 87.8%; Score 36; DB 22; Length 49;

Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DE Protein #4332 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00666.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI, 2001-488899/53.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
XX Claim 15; SEQ ID No 24103; 530pp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC AB21355-AB41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
CC
XX
SQ Sequence 49 AA;
Query Match 87.8%; Score 36; DB 22; Length 49;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGMPLP 6
DB 39 GGMLPH 44
IIIIII:
RESULT 7
ID AAM57746 standard; Protein; 49 AA.
AC AAM57746;
XX
XX 05-NOV-2001 (first entry)
DT
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 29851.
DE
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX

PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI, 2001-483446/52.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 29851; 650pp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX
SQ Sequence 49 AA;
Query Match 87.8%; Score 36; DB 22; Length 49;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGMPLP 6
DB 39 GGMLPH 44
IIIIII:
RESULT 8
ID AAM70159 standard; Protein; 49 AA.
AC AAM70159;
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 30465.
DE
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX
XX WO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 30465; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 49 AA;

Query Match 87.8%; Score 36; DB 22; Length 49;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPPY 6
Db 39 GGWLPH 44

RESULT 9
AAM17988
ID AAM17988 standard; Protein; 49 AA.
XX
AC AAM17988;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #4422 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
OS Homo sapiens.
XX
PN MO200157278-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID NO 22814; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENP; see AAI10066-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 49 AA;

Query Match 87.8%; Score 36; DB 22; Length 49;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPPY 6
Db 39 GGWLPH 44

RESULT 10
AAM30495
ID AAM30495 standard; Protein; 49 AA.
XX
AC AAM30495;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #4532 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN MO200157272-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID NO 30764; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI13315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 49 AA;

Query Match 87.8%; Score 36; DB 22; Length 49;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPPY 6

Db 39 GGMPLH 44

RESULT 11
AAM05626

ID AAM05626 standard; Protein; 49 AA.

AC AAM05626;

DT 09-OCT-2001 (first entry)

DE Peptide #4308 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;

KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

OS Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US000661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-2363599.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-476286/51.

PT Novel single exon nucleic acid probe used to measuring gene expression

PT in a human breast -

XX Claim 27; SEQ ID No 14366; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes

CC (see A100010-A110067). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for measuring human gene expression in

CC a human breast sample, where the probe hybridises at high stringency to a

CC nucleic acid expressed in the human breast. The probes are useful for

CC predicting, diagnosing, grading, staging, monitoring and prognosing

CC diseases of the human breast, particularly those diseases with polygenic

CC aetiology. The diseases include: breast cancer, disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 49 AA:

Query Match 87.8%; Score 36; DB 22; Length 49;

Best Local Similarity 83.3%; Pred. No. 19;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMPLH 6

DB 39 GGMPLH 44

RESULT 12

ABG39797

XX ABG39797 standard; Peptide; 49 AA.

AC ABG39797;

XX 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 29462.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Rudrak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

OS Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-2363599.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 29462; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemodiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocytic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 49 AA:
 Query Match 87.8%; Score 36; DB 23; Length 49;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGWLPPY 6
 Db 39 GGWLPPH 44
 RESULT 13
 AAU71006
 ID AAU71006 standard; Peptide: 20 AA.
 AC AAU71006;
 XX
 DT 25-FEB-2002 (first entry)
 XX
 DE M. tuberculosis Rv0284 protein immunogenic peptide P19.
 XX
 KW Tuberculosis; Tuberculosis; antitubercular; vaccine; Rv0284;
 KW Rv0285; Rv0455c; Rv0556; Rv1386; Rv3477; Rv3878; Rv3879;
 KW M3106.1; ORF13A; Rv0284c; Mycobacterium bovis; Mycobacterium africanum;
 KW BCG vaccine; immunogenic peptide.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200179274-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 19-APR-2001; 2001MO-DK00276.
 XX
 PR 19-APR-2000; 2000DK-0000666.
 PR 21-FEB-2001; 2001DK-0000283.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Agger EM, Andersen P, Okkels LM, Weldingh K;
 XX
 DR WPI; 2002-061970/08.
 XX
 PT New Mycobacterium tuberculosis antigens, useful for diagnosing
 PT tuberculosis, and as a vaccine for treating or preventing infections
 PT caused by species of tuberculosis complex -
 XX
 PS Example 3; Page 95; 11pp; English.
 XX
 CC The invention relates to a substantially pure polypeptide comprising
 CC an amino acid sequence selected from Rv0284, Rv0285, Rv0455c,
 CC Rv0556, Rv1386, Rv3477, Rv3878, Rv3879c or M3106.1 (also
 CC disclosed are ORF13A and Rv0284c), or their immunogenic portion, nucleic
 CC acids encoding them and an amino acid sequence analogue having at least
 CC 70% sequence identity to the polypeptide and is immunogenic.
 CC The protein is useful in preparing a pharmaceutical composition for
 CC diagnosing tuberculosis and in preparing a vaccine against tuberculosis
 CC caused by virulent mycobacteria. The vaccine or immunogenic/
 CC pharmaceutical composition can be used prophylactically in a subject not
 CC infected with a virulent mycobacterium, or therapeutically in a subject
 CC already infected with a virulent mycobacterium. The protein is useful for
 CC preventing, treating and detecting infections caused by species of

CC tuberculosis complex (M. tuberculosis, M. bovis, M. africanum). The
 CC nucleic acids may be used for effecting in vivo expression of the
 CC antigen, and in diagnostic assays for detecting the presence of
 CC pathogenic organisms in a sample. The vaccine is an improvement of the
 CC living BCG vaccine presently available, where one or more copies of the
 CC DNA sequence encoding one or more polypeptide has been incorporated into
 CC the genome of the microorganism to allow the microorganism to express and
 CC secrete the polypeptide. Incorporation of more than one copy of a
 CC nucleotide sequence enhances the immune response. The present
 CC sequence represents an immunogenic peptide derived from an M.
 CC tuberculosis protein of the invention.
 CC
 SQ Sequence 20 AA:
 Query Match 85.4%; Score 35; DB 23; Length 20;
 Best Local Similarity 83.3%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGWLPPY 6
 Db 12 GGWLPPY 17
 RESULT 14
 AA011246
 ID AA011246 standard; Protein: 51 AA.
 XX
 AC AA011246;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 25138.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001MO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AA191177.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 25138; 1399pp + sequence listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polypeptides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 51 AA:

Query Match 85.4%; Score 35; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GMLPY 6
 |||||
 DB 35 GMLPY 39

RESULT 15

ABU11869
 ID ABU11869 standard; Protein; 80 AA.

XX ABU11869;

XX 12-FEB-2003 (first entry)

DE Human secreted protein SECP-24, INCYTE 71728459CD1.

XX Human; SECP; secreted protein; micro-array; liver disease; hepatitis;
 XX cirrhosis; cell proliferative disease; cancer; atherosclerosis;
 XX neurological disorder; epilepsy; Huntington's disease; stroke;
 XX cardiovascular disorder; hypertension; angina pectoris; allergy;
 XX myocardial infarction; immune disorder; inflammatory disorder; AIDS;
 XX hypothyroidism; acquired immunodeficiency syndrome; Cushing's syndrome;
 XX developmental disorder.

XX Homo sapiens.

XX WO200286069-A2.

XX 31-OCT-2002.

XX 19-APR-2002; 2002WO-US12464.

XX 20-APR-2001; 2001US-285207P.
 XX 27-APR-2001; 2001US-287114P.
 XX 03-MAY-2001; 2001US-288640P.
 XX 11-MAY-2001; 2001US-290516P.
 XX 18-MAY-2001; 2001US-292184P.
 XX 21-DEC-2001; 2001US-343553P.
 XX 13-FEB-2002; 2002US-357002P.
 XX 20-FEB-2002; 2002US-358279P.
 XX 19-MAR-2002; 2002US-366041P.

XX (INCY-) INCYTE GENOMICS INC.

XX Klammer AA, Hafalia AJ, Duggan BM, Warren BA, Emerling BM;
 XX Tribouley CM, Arvizu CS, Honchell CD, Nguyen DB, Kallick DA;
 XX Yue H, Au-Young JK, Ramkumar J, Li JK, Thangavelu K, Gietzen KJ;
 XX Ding L, Baughn MR, Yao MG, Walla NK, Mason PM, Lal PG, Graul KC;
 XX Reddy R, Becha SD, Sapperstein SK, Richardson TW, Tran UK;
 XX Elliott VS, Tang YT, Azimzai Y, Yan L, Xu Y;

XX WPI, 2003-093118/08.
 XX N-PSDB; ABX49971.

XX New human secreted proteins (SECP) useful for diagnosing, treating and
 XX preventing diseases or conditions associated with the aberrant SECP
 XX expression e.g. cancer, AIDS, atherosclerosis, epilepsy, allergies,
 XX hepatitis, cirrhosis -

XX Claim 1; Page 161; 192pp; English.

XX The invention relates to an isolated polypeptide comprising any of 30
 XX secreted human proteins (SECP-SECP30) appearing as ABU11846-ABU11875, a
 XX naturally occurring amino acid sequence at least 90-98 % identical to

the sequences, or a biologically active or immunogenic fragment of the
 CC polypeptide. Also included are an isolated polynucleotide encoding SECP,
 CC (including a polynucleotide sequence at least 90-98 % identical to the
 CC sequences, their complements, RNA equivalents or fragments comprising at
 CC least 60 contiguous nucleotides) a recombinant polynucleotide comprising
 CC a promoter sequence operably linked to the SECP polynucleotide, a cell
 CC transformed with the recombinant polynucleotide, an anti-SECP antibody,
 CC comprising the recombinant polynucleotide, an anti-SECP antibody,
 CC screening for antigens of SECP, generating an expression profile of a
 CC sample containing the polynucleotides and an array comprising different
 CC nucleotide molecules affixed at distinct physical locations on a solid
 CC substrate, where at least one nucleotide molecule comprises a first
 CC oligonucleotide or polynucleotide sequence specifically hybridizable with
 CC at least 30 contiguous nucleotides of the target polynucleotide.
 CC The polypeptides and polynucleotides are useful in diagnosing,
 CC treating and preventing diseases or conditions associated with the
 CC decreased expression or overexpression of SECP, such as liver (e.g.
 CC hepatitis, cirrhosis), cell proliferative (e.g. cancer, cancer,
 CC atherosclerosis), neurological (e.g. epilepsy, Huntington's disease,
 CC stroke), cardiovascular (e.g. hypertension, angina pectoris, myocardial
 CC infarction), immune/inflammatory (e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies) and developmental (e.g. Hypothyroidism,
 CC Cushing's syndrome) disorders (many other diseases and conditions
 CC are given in the specification). These are also useful in assessing the
 CC effects of exogenous compounds on the expression of nucleic acid and
 CC amino acid sequences of SECP. The SECP or its fragments are useful in
 CC screening compounds for effectiveness as agonist or antagonist of the
 CC polypeptides, or in altering the expression of the target
 CC polynucleotide and compounds that specifically bind to or modulate the
 CC activity of the polypeptide. The micro-array is useful in monitoring or
 CC measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence is a SECP protein of
 CC the invention.

XX Sequence 80 AA:

Query Match 85.4%; Score 35; DB 24; Length 80;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GMLPY 6
 |||||
 DB 62 GMLPY 67

Search completed: August 20, 2003, 12:33:48
 Job time : 15.1928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 4.15663 seconds

(without alignments)
61.075 Million cell updates/sec

Title: us-09-512-082-33

Perfect score: 41

Sequence: 1 GGMIPY 6

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5a.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5b.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6a.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6b.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCOTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	95.1	497	4	US-09-252-991A-32682
2	35	85.4	104	4	US-09-107-532A-5075
3	35	85.4	220	3	US-09-267-177-8
4	35	85.4	421	4	US-09-198-452A-932
5	35	85.4	505	1	US-09-041-075A-8
6	35	85.4	507	2	US-08-845-161A-4
7	35	85.4	507	3	US-09-270-751-4
8	35	85.4	507	4	US-09-168-218B-5
9	34	82.9	140	4	US-09-252-991A-18847
10	34	82.9	201	4	US-09-252-991A-18847
11	34	82.9	263	4	US-09-252-991A-22895
12	34	82.9	265	4	US-09-252-991A-22211
13	34	82.9	426	4	US-09-252-991A-27172
14	34	82.9	426	4	US-09-252-991A-27172
15	34	82.9	426	4	US-09-252-991A-27172
16	34	82.9	426	4	US-09-252-991A-27172
17	34	82.9	426	4	US-09-252-991A-27172
18	34	82.9	426	4	US-09-252-991A-27172
19	34	82.9	426	4	US-09-252-991A-27172
20	34	82.9	426	4	US-09-252-991A-27172
21	34	82.9	426	4	US-09-252-991A-27172
22	34	82.9	426	4	US-09-252-991A-27172
23	34	82.9	426	4	US-09-252-991A-27172
24	34	82.9	426	4	US-09-252-991A-27172
25	34	82.9	426	4	US-09-252-991A-27172
26	34	82.9	426	4	US-09-252-991A-27172
27	34	82.9	426	4	US-09-252-991A-27172

28	33	80.5	1023	2	US-08-475-891A-2	Sequence 2, Appl1
29	33	80.5	1023	2	US-08-567-375-2	Sequence 2, Appl1
30	33	80.5	1023	2	US-08-587-680A-2	Sequence 2, Appl1
31	33	80.5	1023	2	US-08-567-375-4	Sequence 2, Appl1
32	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
33	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
34	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
35	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
36	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
37	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
38	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
39	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
40	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
41	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
42	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
43	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
44	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1
45	33	80.5	1023	2	US-08-587-680A-4	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-252-991A-32682
; Sequence 32682, Application US/09252991A
; Patient No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196, 136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32682
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32682

Query Match      95.1%; Score 39; DB 4; Length 497;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGMIPY 6
Db      458 GGMIPY 463

RESULT 2
US-09-107-532A-5075
; Sequence 5075, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
```

```
OPERATING SYSTEM: <unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/107,532A
  FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 60/085,598
  FILING DATE: 14 May 1998
  APPLICATION NUMBER: 60/051571
  FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Arinello, Pamela Deneke
  REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (781)893-5007
  TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5075:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 104 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHEICAL: YES
  ORIGINAL SOURCE:
    ORGANISM: Enterococcus faecium
  FEATURE:
    NAME/KEY: misc.feature
    LOCATION: (B) LOCATION 1...104
    SEQUENCE DESCRIPTION: SEQ ID NO: 5075:
US-09-107-532A-5075

Query Match
Best Local Similarity 100.0%; Pred. No. 52; Length 104;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GMLPY 6
      |||||
Db      26 GMLPY 30

RESULT 3
US-09-267-177-8
; Sequence 8, Application US/09267177
; Patent No. 6287856
; GENERAL INFORMATION:
; APPLICANT: Poet, Steven E.
; APPLICANT: Ritchie, Branson W.
; APPLICANT: Niagro, Frank D.
; APPLICANT: Lukert, Phil D.
; TITLE OF INVENTION: Vaccines against Circovirus infections
; FILE REFERENCE: 21099,0057
; CURRENT APPLICATION NUMBER: US/09/267,177
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 60/077,890
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: beak and feather disease virus
US-09-267-177-8

Query Match
Best Local Similarity 85.4%; Score 35; DB 3; Length 220;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GMLPY 6
      |||||
Db      138 GMLPY 142
```

```
RESULT 4
US-09-198-452A-932
; Sequence 932, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 932
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-932

Query Match
Best Local Similarity 85.4%; Score 35; DB 4; Length 421;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GMLPY 6
      |||||
Db      308 GMLPY 312

RESULT 5
US-09-041-075A-8
; Sequence 8, Application US/09041075A
; Patent No. H002022
; GENERAL INFORMATION:
; APPLICANT: Heidler, Steven A
; APPLICANT: Radding, Jeffrey A
; TITLE OF INVENTION: IPC SYNTHASE GENES FROM FUNGI
; FILE REFERENCE: X-11242 Sequence 1st
; Patent No. H002022
; CURRENT APPLICATION NUMBER: US/09/041,075A
; CURRENT FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/043,591
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Candida parapsilosis
US-09-041-075A-8

Query Match
Best Local Similarity 85.4%; Score 35; DB 1; Length 505;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GMLPY 6
      |||||
Db      165 GMLPY 169

RESULT 6
US-08-845-161A-4
; Sequence 4, Application US/08845161A
; Patent No. 5976850
; GENERAL INFORMATION:
; APPLICANT: Latche, Richard
; APPLICANT: Rose, Kenneth A.
; APPLICANT: Stapleton, Genevieve
; TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA SEQUENCES CODING THEREFOR AND USED THEREOF
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NO. 5976850th Glebe Rd. 8th floor
```

CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,161A
FILING DATE: 21-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02465
FILING DATE: 18-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9421093.7
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 604-408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-845-161A-4

Query Match
Best Local Similarity 85.4%; Score 35; DB 2; Length 507;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 49 GWLPY 53

RESULT 7
US-09-270-751-4
Sequence 4, Application US/09270751
Patent No. 6184350
GENERAL INFORMATION:
APPLICANT: Lathe, Richard
Rose, Kenneth A.
Stapleton, Genevieve
TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS: DNA
SEQUENCES CODING THEREFOR AND USED THEREOF
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 6184350th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,751
FILING DATE: 17-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02465
FILING DATE: 18-OCT-1995

APPLICATION NUMBER: GB 9421093.7
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 604-408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-270-751-4

Query Match
Best Local Similarity 85.4%; Score 35; DB 3; Length 507;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 49 GWLPY 53

RESULT 8
US-09-168-218B-5
Sequence 5, Application US/09168218B
Patent No. 6420353
GENERAL INFORMATION:
APPLICANT: LATHE, RICHARD F.
ROSE, KENNETH A.
APPLICANT: SECKL, JOHNATHAN R.
APPLICANT: BEST, RUTH
APPLICANT: YAU, JOYCE L.W.
APPLICANT: LECKIE, CAROLINE M.
TITLE OF INVENTION: NEUROSTEROIDS
FILE REFERENCE: 604-460
CURRENT APPLICATION NUMBER: US/09/168,218B
CURRENT FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 9607289.7
PRIOR FILING DATE: 1996-04-09
PRIOR APPLICATION NUMBER: 9608445.5
PRIOR FILING DATE: 1996-04-24
PRIOR APPLICATION NUMBER: 9704905.0
PRIOR FILING DATE: 1997-03-10
PRIOR APPLICATION NUMBER: PCT/GB97/00955
PRIOR FILING DATE: 1997-04-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 507
TYPE: PRT
ORGANISM: Murine sp.
US-09-168-218B-5

Query Match
Best Local Similarity 85.4%; Score 35; DB 4; Length 507;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 49 GWLPY 53

RESULT 9
US-09-252-991A-18847
Sequence 18847, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 18847
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18847

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLP 5
11111
DB 122 GGMLP 126

RESULT 10
US-09-252-991A-29895
; Sequence 29895, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29895
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29895

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLP 5
11111
DB 18 GGMLP 22

RESULT 11
US-09-252-991A-22211
; Sequence 22211, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22211

;; LENGTH: 263
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: (263)
;; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-22211

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLP 5
11111
DB 233 GGMLP 237

RESULT 12
US-09-252-991A-27172
; Sequence 27172, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27172
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27172

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLP 5
11111
DB 254 GGMLP 258

RESULT 13
US-09-599-360B-106
; Sequence 106, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouquelerey, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 106
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo Sapiens

FEATURE:
NAME/KEY: SIGNAL
LOCATION: -184...-1
US-09-599-360B-106

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 385;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLP 5
|||||
Db 171 GGWLP 175

RESULT 14
US-09-268-311-2
Sequence 2, Application US/09268311
Patent No. 6482923
GENERAL INFORMATION:
APPLICANT: Shi, Yangu
TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
FILE REFERENCE: PF398P1
CURRENT APPLICATION NUMBER: US/09/268,311
CURRENT FILING DATE: 1999-03-16
EARLIER APPLICATION NUMBER: 60/059,133
EARLIER FILING DATE: 1997-09-17
EARLIER APPLICATION NUMBER: 09/154,219
EARLIER FILING DATE: 1998-09-16
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-268-311-2

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 426;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLP 5
|||||
Db 288 GGWLP 292

RESULT 15
US-09-268-311-3
Sequence 3, Application US/09268311
Patent No. 6482923
GENERAL INFORMATION:
APPLICANT: Shi, Yangu
APPLICANT: Ruben, Steve M.
TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
FILE REFERENCE: PF398P1
CURRENT APPLICATION NUMBER: US/09/268,311
CURRENT FILING DATE: 1999-03-16
EARLIER APPLICATION NUMBER: 60/059,133
EARLIER FILING DATE: 1997-09-17
EARLIER APPLICATION NUMBER: 09/154,219
EARLIER FILING DATE: 1998-09-16
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-268-311-3

Query Match
Best Local Similarity 82.9%; Score 34; DB 4; Length 426;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLP 5
|||||
Db 288 GGWLP 292

Search completed: August 20, 2003, 12:44:23
Job time : 5.15663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 6.50602 Seconds
(without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GCMWLPY 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications, AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	6	11 US-09-300-425B-33	Sequence 33, Appl
2	37	90.2	51	10 US-09-989-919-85	Sequence 85, Appl
3	36	87.8	49	9 US-09-864-761-37631	Sequence 37631, A
4	35	85.4	55	15 US-10-083-357-859	Sequence 859, App
5	35	85.4	61	15 US-10-083-357-997	Sequence 907, App
6	35	85.4	61	15 US-10-083-357-934	Sequence 934, App
7	35	85.4	237	10 US-09-853-450-32	Sequence 32, Appl
8	35	85.4	246	15 US-10-156-761-13728	Sequence 13728, A
9	35	85.4	251	16 US-10-278-536-196	Sequence 196, App
10	35	85.4	505	9 US-09-742-582-8	Sequence 8, Appl1
11	35	85.4	505	11 US-09-742-580-8	Sequence 8, Appl1
12	35	85.4	505	11 US-09-742-581-8	Sequence 8, Appl1
13	35	85.4	505	15 US-10-156-761-9541	Sequence 9541, Ap
14	35	85.4	520	15 US-10-156-761-12623	Sequence 12623, A
15	34	82.9	98	9 US-09-864-761-41798	Sequence 41798, A

16	34	82.9	145	10	US-09-764-868-1233	Sequence 1233, Ap
17	34	82.9 <td>156 <td>10 <td>US-09-981-876-136</td> <td>Sequence 136, App</td> </td></td>	156 <td>10 <td>US-09-981-876-136</td> <td>Sequence 136, App</td> </td>	10 <td>US-09-981-876-136</td> <td>Sequence 136, App</td>	US-09-981-876-136	Sequence 136, App
18	34	82.9 <td>156 <td>11 <td>US-09-148-545-136</td> <td>Sequence 136, App</td> </td></td>	156 <td>11 <td>US-09-148-545-136</td> <td>Sequence 136, App</td> </td>	11 <td>US-09-148-545-136</td> <td>Sequence 136, App</td>	US-09-148-545-136	Sequence 136, App
19	34	82.9 <td>197 <td>12 <td>US-10-017-161-1600</td> <td>Sequence 1600, Ap</td> </td></td>	197 <td>12 <td>US-10-017-161-1600</td> <td>Sequence 1600, Ap</td> </td>	12 <td>US-10-017-161-1600</td> <td>Sequence 1600, Ap</td>	US-10-017-161-1600	Sequence 1600, Ap
20	34	82.9 <td>214 <td>10 <td>US-09-738-626-4691</td> <td>Sequence 4691, Ap</td> </td></td>	214 <td>10 <td>US-09-738-626-4691</td> <td>Sequence 4691, Ap</td> </td>	10 <td>US-09-738-626-4691</td> <td>Sequence 4691, Ap</td>	US-09-738-626-4691	Sequence 4691, Ap
21	34	82.9 <td>289</td> <td>10 <td>US-09-778-971-8</td> <td>Sequence 8, Appl1</td> </td>	289	10 <td>US-09-778-971-8</td> <td>Sequence 8, Appl1</td>	US-09-778-971-8	Sequence 8, Appl1
22	34	82.9 <td>314</td> <td>15 <td>US-10-219-834-79</td> <td>Sequence 79, Appl1</td> </td>	314	15 <td>US-10-219-834-79</td> <td>Sequence 79, Appl1</td>	US-10-219-834-79	Sequence 79, Appl1
23	34	82.9 <td>322</td> <td>12 <td>US-09-995-225-20</td> <td>Sequence 20, Appl1</td> </td>	322	12 <td>US-09-995-225-20</td> <td>Sequence 20, Appl1</td>	US-09-995-225-20	Sequence 20, Appl1
24	34	82.9 <td>322</td> <td>10 <td>US-09-995-225-20</td> <td>Sequence 20, Appl1</td> </td>	322	10 <td>US-09-995-225-20</td> <td>Sequence 20, Appl1</td>	US-09-995-225-20	Sequence 20, Appl1
25	34	82.9 <td>322</td> <td>15 <td>US-10-183-116-31</td> <td>Sequence 31, Appl1</td> </td>	322	15 <td>US-10-183-116-31</td> <td>Sequence 31, Appl1</td>	US-10-183-116-31	Sequence 31, Appl1
26	34	82.9 <td>332</td> <td>15 <td>US-10-225-567A-674</td> <td>Sequence 674, App</td> </td>	332	15 <td>US-10-225-567A-674</td> <td>Sequence 674, App</td>	US-10-225-567A-674	Sequence 674, App
27	34	82.9 <td>328</td> <td>15 <td>US-10-219-834-20</td> <td>Sequence 20, Appl1</td> </td>	328	15 <td>US-10-219-834-20</td> <td>Sequence 20, Appl1</td>	US-10-219-834-20	Sequence 20, Appl1
28	34	82.9 <td>354</td> <td>10 <td>US-09-764-864-824</td> <td>Sequence 824, App</td> </td>	354	10 <td>US-09-764-864-824</td> <td>Sequence 824, App</td>	US-09-764-864-824	Sequence 824, App
29	34	82.9 <td>371</td> <td>10 <td>US-09-981-876-259</td> <td>Sequence 259, App</td> </td>	371	10 <td>US-09-981-876-259</td> <td>Sequence 259, App</td>	US-09-981-876-259	Sequence 259, App
30	34	82.9 <td>371</td> <td>11 <td>US-09-148-545-259</td> <td>Sequence 259, App</td> </td>	371	11 <td>US-09-148-545-259</td> <td>Sequence 259, App</td>	US-09-148-545-259	Sequence 259, App
31	34	82.9 <td>350</td> <td>12 <td>US-10-183-708-102</td> <td>Sequence 102, App</td> </td>	350	12 <td>US-10-183-708-102</td> <td>Sequence 102, App</td>	US-10-183-708-102	Sequence 102, App
32	34	82.9 <td>499</td> <td>10 <td>US-09-778-971-7</td> <td>Sequence 7, Appl1</td> </td>	499	10 <td>US-09-778-971-7</td> <td>Sequence 7, Appl1</td>	US-09-778-971-7	Sequence 7, Appl1
33	34	82.9 <td>502</td> <td>9 <td>US-09-886-404-18</td> <td>Sequence 18, Appl1</td> </td>	502	9 <td>US-09-886-404-18</td> <td>Sequence 18, Appl1</td>	US-09-886-404-18	Sequence 18, Appl1
34	34	82.9 <td>502</td> <td>10 <td>US-09-778-971-5</td> <td>Sequence 5, Appl1</td> </td>	502	10 <td>US-09-778-971-5</td> <td>Sequence 5, Appl1</td>	US-09-778-971-5	Sequence 5, Appl1
35	34	82.9 <td>502</td> <td>10 <td>US-09-874-503-12</td> <td>Sequence 12, Appl1</td> </td>	502	10 <td>US-09-874-503-12</td> <td>Sequence 12, Appl1</td>	US-09-874-503-12	Sequence 12, Appl1
36	34	82.9 <td>502</td> <td>11 <td>US-09-816-744-12</td> <td>Sequence 12, Appl1</td> </td>	502	11 <td>US-09-816-744-12</td> <td>Sequence 12, Appl1</td>	US-09-816-744-12	Sequence 12, Appl1
37	34	82.9 <td>502</td> <td>11 <td>US-09-747-259-12</td> <td>Sequence 12, Appl1</td> </td>	502	11 <td>US-09-747-259-12</td> <td>Sequence 12, Appl1</td>	US-09-747-259-12	Sequence 12, Appl1
38	34	82.9 <td>502</td> <td>11 <td>US-09-908-827-12</td> <td>Sequence 12, Appl1</td> </td>	502	11 <td>US-09-908-827-12</td> <td>Sequence 12, Appl1</td>	US-09-908-827-12	Sequence 12, Appl1
39	34	82.9 <td>502</td> <td>11 <td>US-09-863-818A-2</td> <td>Sequence 2, Appl1</td> </td>	502	11 <td>US-09-863-818A-2</td> <td>Sequence 2, Appl1</td>	US-09-863-818A-2	Sequence 2, Appl1
40	34	82.9 <td>502</td> <td>12 <td>US-10-063-735-158</td> <td>Sequence 158, App</td> </td>	502	12 <td>US-10-063-735-158</td> <td>Sequence 158, App</td>	US-10-063-735-158	Sequence 158, App
41	34	82.9 <td>502</td> <td>12 <td>US-10-199-672-400</td> <td>Sequence 400, App</td> </td>	502	12 <td>US-10-199-672-400</td> <td>Sequence 400, App</td>	US-10-199-672-400	Sequence 400, App
42	34	82.9 <td>502</td> <td>12 <td>US-10-187-748-400</td> <td>Sequence 400, App</td> </td>	502	12 <td>US-10-187-748-400</td> <td>Sequence 400, App</td>	US-10-187-748-400	Sequence 400, App
43	34	82.9 <td>502</td> <td>12 <td>US-10-194-457-400</td> <td>Sequence 400, App</td> </td>	502	12 <td>US-10-194-457-400</td> <td>Sequence 400, App</td>	US-10-194-457-400	Sequence 400, App
44	34	82.9 <td>502</td> <td>12 <td>US-10-006-867-158</td> <td>Sequence 158, App</td> </td>	502	12 <td>US-10-006-867-158</td> <td>Sequence 158, App</td>	US-10-006-867-158	Sequence 158, App
45	34	82.9 <td>502</td> <td>14 <td>US-10-052-586-400</td> <td>Sequence 400, App</td> </td>	502	14 <td>US-10-052-586-400</td> <td>Sequence 400, App</td>	US-10-052-586-400	Sequence 400, App

ALIGNMENTS

RESULT 1
US-09-300-425B-33
; Sequence 33, Application US/09300425B
; Publication No. US20030045681A1
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITTI, Francesca
; APPLICANT: BIRCHIER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
; OTHER INFORMATION: antibody clone
US-09-300-425B-33

Query Match 100.0%; Score 41; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCMWLPY 6
DB 1 GCMWLPY 6

RESULT 2

US-09-989-919-85
; Sequence 85, Application US/09989919
; Patent No. US20020164344A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Pluta, Jason
; APPLICANT: Ghosh, Malavika
; APPLICANT: Sun, Yonngang
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro
; FILE REFERENCE: DEX-0289
; CURRENT APPLICATION NUMBER: US/09/989,919
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,505
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-919-85
Query Match 90.2%; Score 37; DB 10; Length 51;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GMLPY 6
|||||
DB 43 GMLPF 48
RESULT 3
US-09-864-761-37631
; Sequence 37631, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37631
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO A1229041.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 15
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 21
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 22
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 17
; OTHER INFORMATION: EST_HUMAN HIT: AW468873.1, EVALUE 2.00e-01
US-09-864-761-37631
Query Match 87.8%; Score 36; DB 9; Length 49;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GMLPY 6
|||||
DB 39 GMLPH 44
RESULT 4
US-10-083-357-859
; Sequence 859, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Olandong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 859
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-859
Query Match 85.4%; Score 35; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GMLPY 6
|||||
DB 2 GMLPY 6
RESULT 5
US-10-083-357-907
; Sequence 907, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Olandong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes

FILE REFERENCE: 032796-090
CURRENT APPLICATION NUMBER: US/10/083,357
CURRENT FILING DATE: 2002-02-27
NUMBER OF SEQ ID NOS: 1346
SEQ ID NO 907
LENGTH: 61
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-083-357-907

Query Match
Best Local Similarity 85.4%; Score 35; DB 15; Length 61;
Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 2 GWLPY 6

RESULT 6
US-10-083-357-934
Sequence 934, Application US/10083357
Publication No. US20030054370A1
GENERAL INFORMATION:
APPLICANT: Qiantong Zeng et al.
TITLE OF INVENTION: Systemic Discovery of New Genes
FILE REFERENCE: 032796-090
CURRENT APPLICATION NUMBER: US/10/083,357
CURRENT FILING DATE: 2002-02-27
NUMBER OF SEQ ID NOS: 1346
SEQ ID NO 934
LENGTH: 61
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-083-357-934

Query Match
Best Local Similarity 85.4%; Score 35; DB 15; Length 61;
Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 2 GWLPY 6

RESULT 7
US-09-853-450-32
Sequence 32, Application US/09853450
Publication No. US20020194645A1
GENERAL INFORMATION:
APPLICANT: Yanofsky, Martin F.
APPLICANT: Pelaz, Soraya
APPLICANT: Ditta, Gary
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
FILE REFERENCE: 19452A-002400US
CURRENT APPLICATION NUMBER: US/09/853,450
CURRENT FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 237
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: SEPALATA3 (SEP3)
US-09-853-450-32

Query Match
Best Local Similarity 85.4%; Score 35; DB 10; Length 237;
Pred. No. 2,7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 228 GWLPY 232

RESULT 8
US-10-156-761-13728
Sequence 13728, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADATOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13728
LENGTH: 246
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-13728

Query Match
Best Local Similarity 85.4%; Score 35; DB 15; Length 246;
Pred. No. 2,8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
Db 69 GWLPY 73

RESULT 9
US-10-278-536-196
Sequence 196, Application US/10278536
Publication No. US2003013186A1
GENERAL INFORMATION:
APPLICANT: Samaha, Raymond
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Keddie, James
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Adam, Luc
APPLICANT: Brown, Pierre
TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-011
CURRENT APPLICATION NUMBER: US/10/278,536
CURRENT FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 196
LENGTH: 251
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G145
US-10-278-536-196

Query Match 85.4%; Score 35; DB 16; Length 251;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 |||||
 Db 242 GWLPY 246

RESULT 10
 US-09-742-582-8
 ; Sequence 8, Application US/09742582
 ; Patent No. US20010012615A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heidler, Steven A
 ; APPLICANT: Radding, Jeffrey A
 ; TITLE OF INVENTION: IPC SYNTHASE GENES FROM FUNGI
 ; FILE REFERENCE: X-11242 Sequence 1st
 ; Patent No. US20010012615A1
 ; CURRENT APPLICATION NUMBER: US/09/742,582
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/041,075
 ; PRIOR FILING DATE: 1998-03-10
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 505
 ; TYPE: PRF
 ; ORGANISM: Candida parapsilosis
 US-09-742-582-8

Query Match 85.4%; Score 35; DB 9; Length 505;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 |||||
 Db 165 GWLPY 169

RESULT 11
 US-09-742-580-8
 ; Sequence 8, Application US/09742580
 ; Publication No. US20030022342A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heidler, Steven A
 ; APPLICANT: Radding, Jeffrey A
 ; TITLE OF INVENTION: IPC SYNTHASE GENES FROM FUNGI
 ; FILE REFERENCE: X-11242 Sequence 1st
 ; CURRENT APPLICATION NUMBER: US/09/742,580
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/041,075
 ; PRIOR FILING DATE: 1998-03-10
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 505
 ; TYPE: PRF
 ; ORGANISM: Candida parapsilosis
 US-09-742-580-8

Query Match 85.4%; Score 35; DB 11; Length 505;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 |||||
 Db 165 GWLPY 169

RESULT 12
 US-09-742-581-8

; Sequence 8, Application US/09742581
 ; Publication No. US20030022343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Heidler, Steven A
 ; APPLICANT: Radding, Jeffrey A
 ; TITLE OF INVENTION: IPC SYNTHASE GENES FROM FUNGI
 ; FILE REFERENCE: X-11242 Sequence 1st
 ; CURRENT APPLICATION NUMBER: US/09/742,581
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/041,075
 ; PRIOR FILING DATE: 1998-03-10
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 505
 ; TYPE: PRF
 ; ORGANISM: Candida parapsilosis
 US-09-742-581-8

Query Match 85.4%; Score 35; DB 11; Length 505;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 |||||
 Db 165 GWLPY 169

RESULT 13
 US-10-156-761-9541
 ; Sequence 9541, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 9541
 ; LENGTH: 505
 ; TYPE: PRF
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-9541

Query Match 85.4%; Score 35; DB 15; Length 505;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 |||||
 Db 159 GWLPY 163

RESULT 14
 US-10-156-761-12623
 ; Sequence 12623, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI

```

; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12623
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12623

Query Match      85.4%; Score 35; DB 15; Length 520;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GGLPYP 6
        |||:|
Db      183 GGMPY 188

RESULT 15
US-09-864-761-41798
; Sequence 41798, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
```

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41798
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC020557.2
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL - 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.8
; OTHER INFORMATION: EST_HUMAN HIT: D54392.1, EVALUATE 4.60e+00
; OTHER INFORMATION: SWISSPROT HIT: Q98910, EVALUATE 1.10e+00
US-09-864-761-41798

Query Match      82.9%; Score 34; DB 9; Length 98;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGLPYP 6
        |||:|
Db      36 GGMPW 41
```

Search completed: August 20, 2003, 13:16:47
Job time : 6.50602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 12:35:06 ; Search time 61.4096 Seconds
(Without alignments)
85.038 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GWWLPY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	6	14	US-09-075-338C-33
2	41	100.0	6	17	US-09-300-425B-33

3	41	100.0	6	19	US-09-512-082-33	Sequence 33, Appl
4	41	100.0	238	22	US-09-791-537-78364	Sequence 78364, A
5	39	95.1	457	30	US-10-419-128-32682	Sequence 32682, A
6	38	92.7	107	30	US-10-437-963-112199	Sequence 112199, A
7	37	90.2	51	1	PCT-US02-04240-67	Sequence 67, Appl
8	37	90.2	51	25	US-09-989-919-85	Sequence 85, Appl
9	37	90.2	150	22	US-09-760-468-1397	Sequence 1397, Ap
10	37	90.2	150	28	US-10-216-568-1397	Sequence 1397, Ap
11	37	90.2	115	30	US-10-437-963-104557	Sequence 104557, A
12	37	90.2	211	30	US-10-424-599-11560	Sequence 11560, A
13	37	90.2	382	28	US-10-219-999-33023	Sequence 33023, A
14	37	90.2	382	30	US-10-425-114-39507	Sequence 39507, A
15	37	90.2	382	31	US-60-312-544-5701	Sequence 5701, A
16	37	90.2	714	30	US-10-437-963-158571	Sequence 158571, A
17	36	87.8	49	1	PCT-US01-00663-30764	Sequence 30764, A
18	36	87.8	49	23	US-09-864-761-37631	Sequence 37631, A
19	36	87.8	49	27	US-10-182-993-29851	Sequence 29851, A
20	36	87.8	49	27	US-10-182-993-24103	Sequence 24103, A
21	36	87.8	49	27	US-10-182-997-22814	Sequence 22814, A
22	36	87.8	49	27	US-10-182-998-14366	Sequence 14366, A
23	36	87.8	49	27	US-10-203-134-30465	Sequence 30465, A
24	36	87.8	49	28	US-10-203-136-29462	Sequence 29462, A
25	36	87.8	49	28	US-10-203-137-30764	Sequence 30764, A
26	36	87.8	49	28	US-10-203-138-14758	Sequence 14758, A
27	36	87.8	49	28	US-10-203-138A-14758	Sequence 14758, A
28	36	87.8	49	28	US-10-203-138-14758	Sequence 14758, A
29	36	87.8	49	28	US-10-203-138-14758	Sequence 14758, A
30	36	87.8	49	31	US-60-236-359-19952	Sequence 19952, A
31	36	87.8	216	30	US-10-437-963-158898	Sequence 158898, A
32	36	87.8	234	30	US-10-424-599-191656	Sequence 191656, A
33	36	87.8	596	22	US-09-791-537-66033	Sequence 66033, A
34	35	85.4	20	22	US-10-274-207-72	Sequence 72, Appl
35	35	85.4	55	26	US-10-083-357-859	Sequence 859, Appl
36	35	85.4	55	31	US-60-333-726-839	Sequence 839, Appl
37	35	85.4	58	26	US-10-099-056-1423	Sequence 1423, Ap
38	35	85.4	61	26	US-10-083-357-907	Sequence 907, Appl
39	35	85.4	61	26	US-10-083-357-934	Sequence 934, Appl
40	35	85.4	61	31	US-60-333-726-907	Sequence 907, Appl
41	35	85.4	61	31	US-60-333-726-934	Sequence 934, Appl
42	35	85.4	80	1	PCT-US02-12464-74	Sequence 74, Appl
43	35	85.4	85	30	US-10-437-963-194480	Sequence 194480, A
44	35	85.4	104	30	US-10-417-884-5075	Sequence 5075, Ap
45	35	85.4	131	22	US-09-758-471-3759	Sequence 3759, Ap

ALIGNMENTS

RESULT 1
US-09-075-338C-33
Sequence 33, Application US/09075338C
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARLI, Lorenzo
APPLICANT: VITTI, Francesca
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
FILE REFERENCE: SCH-1733
CURRENT APPLICATION NUMBER: US/09/075,338C
CURRENT FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
OTHER INFORMATION: antibody clone
US-09-075-338C-33

Query Match 100.0% Score 41, DB 14, Length 6;
Best Local Similarity 100.0%; Pred. No. 5.2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
Db 1 GGMLPY 6

RESULT 2

US-09-300-425B-33
; Sequence 33, Application US/09300425B
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARI, Lorenzo
; APPLICANT: VITI, Francesca
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 33
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-300-425B-33

Query Match
Best Local Similarity 100.0%; Score 41; DB 17; Length 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
Db 1 GGMLPY 6

RESULT 3

US-09-512-082-33
; Sequence 33, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; FILE REFERENCE: SCH-1733P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 33
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-512-082-33

Query Match 100.0%; Score 41; DB 19; Length 6;

Best Local Similarity 100.0%; Pred. No. 5.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
Db 1 GGMLPY 6

RESULT 4

US-09-791-537-78364
; Sequence 78364, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 78364
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-78364

Query Match
Best Local Similarity 100.0%; Score 41; DB 22; Length 238;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
Db 222 GGMLPY 227

RESULT 5

US-10-419-128-32682
; Sequence 32682, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 32682
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-32682

Query Match
Best Local Similarity 95.1%; Score 39; DB 30; Length 497;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
Db 458 GGMLPY 463

RESULT 6

US-10-437-963-112199
; Sequence 112199, Application US/10437963
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 112199
LENGTH: 107
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_16106C.1.pep
US-10-437-963-112199

Query Match
Best Local Similarity 92.7%; Score 38; DB 30; Length 107;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
DB 11 GGMLPY 16

RESULT 7
PCT-US02-04240-67
Sequence 67, Application PC/TUS0204240
GENERAL INFORMATION:
APPLICANT: diadexus, Inc.
APPLICANT: Pideric, Alejandra
APPLICANT: Macina, Roberto
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Pro
FILE REFERENCE: DEX-0309
CURRENT APPLICATION NUMBER: PCT/US02/04240
CURRENT FILING DATE: 2002-02-12
PRIOR APPLICATION NUMBER: 60/268,291
PRIOR FILING DATE: 2001-02-13
NUMBER OF SEQ ID NOS: 91
SOFTWARE: PatentIn version 3.1
SEQ ID NO 67
LENGTH: 51
TYPE: PRT
ORGANISM: Homo sapien
PCT-US02-04240-67

Query Match
Best Local Similarity 90.2%; Score 37; DB 1; Length 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
DB 43 GGMLPY 48

RESULT 8
US-09-989-919-85
Sequence 85, Application US/09989919
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Heve
APPLICANT: Pluta, Jason
APPLICANT: Ghosh, Malavika
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
FILE REFERENCE: DEX-0289
CURRENT APPLICATION NUMBER: US/09/989,919
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,505
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 124
SOFTWARE: PatentIn version 3.1
SEQ ID NO 85
LENGTH: 51
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-919-85

Query Match
Best Local Similarity 90.2%; Score 37; DB 25; Length 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
DB 43 GGMLPY 48

RESULT 9
US-09-760-469-1397
Sequence 1397, Application US/09760469
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT254
CURRENT APPLICATION NUMBER: US/09/760,469
CURRENT FILING DATE: 2001-01-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1983
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1397
LENGTH: 150
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (39)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (92)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (110)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (115)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (117)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-469-1397

Query Match
Best Local Similarity 90.2%; Score 37; DB 22; Length 150;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMLPY 6
|||||
DB 140 GGMLPY 145

RESULT 10

```

US-10-216-583-1397
; Sequence 1397, Application US/10216583
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT254C1N
; CURRENT APPLICATION NUMBER: US/10/216,583
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/760,469
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1983
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1397
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (39)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (92)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (148)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-216-583-1397
Query Match          90.2%; Score 37; DB 28; Length 150;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGWLPY 6

DB 140 GGWLPF 145

```

RESULT 11
US-10-437-963-104557
; Sequence 104557, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104557
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(175)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101880C.1.pep
US-10-437-963-104557
Query Match          90.2%; Score 37; DB 30; Length 175;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGWLPY 6

DB 170 GGWLPF 175

```

RESULT 12
US-10-424-599-171560
; Sequence 171560, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 171560
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125934C.1.pep
US-10-424-599-171560
Query Match          90.2%; Score 37; DB 30; Length 211;
Best Local Similarity 83.3%; Pred. No. 9.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGWLPY 6

DB 169 GGWLPF 174

RESULT 13
US-10-219-999-33023

```
Sequence 33023, Application US/10219999
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10(52726)C
CURRENT APPLICATION NUMBER: US/10/219,999
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 60/324,109
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/312,544
PRIOR FILING DATE: 2001-08-15
NUMBER OF SEQ ID NOS: 63520
SEQ ID NO 33023
LENGTH: 382
TYPE: PRT
ORGANISM: Zea mays
US-10-219-999-33023
```

```
Query Match
Best Local Similarity 90.2%; Score 37; DB 28; Length 382;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGWLPY 6
DB 163 GGWLPF 168
```

```
RESULT 14
US-10-425-114-39507
Sequence 39507, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 39507
LENGTH: 382
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700238067_FLI.pep
US-10-425-114-39507
```

```
Query Match
Best Local Similarity 90.2%; Score 37; DB 30; Length 382;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGWLPY 6
DB 163 GGWLPF 168
```

```
RESULT 15
US-60-312-544-5701
Sequence 5701, Application US/60312544
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10(52726)A
CURRENT APPLICATION NUMBER: US/60/312,544
CURRENT FILING DATE: 2001-08-15
NUMBER OF SEQ ID NOS: 10730
SEQ ID NO 5701
LENGTH: 382
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700238067_FLI
US-60-312-544-5701
```

```
Query Match
Best Local Similarity 90.2%; Score 37; DB 31; Length 382;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GGWLPY 6
DB 163 GGWLPF 168
```

```
Search completed: August 20, 2003, 13:13:45
Job time : 63.4096 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:34:01 ; Search time 1.73494 Seconds

(without alignments)
91.7710 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GGWLPY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	82.9	131	6	US-10-294-433-253
2	34	82.9	294	6	US-10-286-897-1865
3	34	82.9	294	6	US-10-258-898A-1865
4	34	82.9	310	6	US-10-286-897-5437
5	34	82.9	310	6	US-10-258-898A-5437
6	34	82.9	312	1	PCT-US03-19871-10
7	34	82.9	312	6	US-10-408-765A-1125
8	34	82.9	322	6	US-10-401-397A-2
9	34	82.9	322	6	US-10-292-798-1274
10	34	82.9	328	1	PCT-US03-19871-8
11	34	82.9	371	1	PCT-US03-19871-6
12	34	82.9	416	6	US-10-603-114-5455
13	34	82.9	486	1	PCT-US03-19871-2
14	34	82.9	502	6	US-10-617-573-12
15	34	82.9	531	1	PCT-US03-19871-4
16	34	82.9	589	6	US-10-603-114-6420
17	34	82.9	628	6	US-10-627-476-550
18	34	82.9	628	6	US-10-627-476-666
19	34	82.9	647	6	US-10-603-113-26574
20	33	80.5	361	6	US-10-009-330C-5
21	33	80.5	498	6	US-10-603-113-15018
22	33	80.5	567	6	US-10-603-114-7515
23	32	78.0	80	6	US-10-291-172-265
24	32	78.0	98	6	US-10-612-783-4322
25	32	78.0	188	6	US-10-603-114-6230
26	32	78.0	210	6	US-10-603-113-14700

27	32	78.0	803	6	US-10-603-114-4623	Sequence 4623, Ap
28	31	75.6	239	6	US-10-603-113-15999	Sequence 15999, A
29	31	75.6	361	6	US-10-460-614-50	Sequence 50, Appl
30	31	75.6	364	6	US-10-603-114-7525	Sequence 7525, Ap
31	31	75.6	404	6	US-10-603-114-7864	Sequence 7864, Ap
32	31	75.6	499	1	PCT-US02-18638A-26	Sequence 26, Appl
33	31	75.6	511	6	US-10-612-783-3649	Sequence 3649, Ap
34	31	75.6	599	6	US-10-612-783-5154	Sequence 5154, Ap
35	31	75.6	638	1	PCT-US02-41847A-11	Sequence 11, Appl
36	31	75.6	786	6	US-10-603-114-6379	Sequence 6379, Ap
37	31	75.6	822	1	PCT-US02-41847A-6	Sequence 6, Appl1
38	31	75.6	822	1	PCT-US02-41847A-6	Sequence 9338, Ap
39	31	75.6	1074	6	US-10-273-573-9338	Sequence 41, Appl
40	31	75.6	1074	6	US-10-369-072-41	Sequence 44, Appl
41	31	75.6	1074	6	US-10-369-072-44	Sequence 43, Appl
42	31	75.6	1077	6	US-10-369-072-43	Sequence 256, Ap
43	31	75.6	1372	6	US-10-294-433-256	Sequence 688, Ap
44	31	75.6	1380	6	US-10-294-433-688	Sequence 1625, Ap
45	31	75.6	1938	7	US-60-490-890-1625	

ALIGNMENTS

RESULT 1
US-10-294-433-253
; Sequence 253, Application US/10294433
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 792CIP4
; CURRENT APPLICATION NUMBER: US/10/294,433
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/14826
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/989,600
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/677,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 10/150,802
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/715,869
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 10/167,379
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/775,330
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 253
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-433-253

Query Match 82.9%; Score 34; DB 6; Length 131;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GGWLP 5
70 GGWLP 74

RESULT 2
US-10-286-897-1865
; Sequence 1865, Application US/10286897

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_fl_genes_b Versions 1.0
; SEQ ID NO 1865
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-1865
```

```

Query Match      82.9%; Score 34; DB 6; Length 294;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGMLP 5
        |||||
Db       7 GGMLP 11
```

```

RESULT 3
US-10-258-898A-1865
; Sequence 1865, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_fl_genes_b Versions 1.0
; SEQ ID NO 1865
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-1865
```

```
Query Match      82.9%; Score 34; DB 6; Length 294;
```

```

Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGMLP 5
        |||||
Db       7 GGMLP 11
```

```

RESULT 4
US-10-286-897-5437
; Sequence 5437, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_fl_genes_b Versions 1.0
; SEQ ID NO 5437
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-5437
```

```

Query Match      82.9%; Score 34; DB 6; Length 310;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGMLP 5
        |||||
Db       23 GGMLP 27
```

```

RESULT 5
US-10-258-898A-5437
; Sequence 5437, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
```

PRIOR APPLICATION NUMBER: US09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7143
SOFTWARE: PL_FL_genes_b Versions 1.0
SEQ ID NO 5437
LENGTH: 310
TYPE: PRT
ORGANISM: Homo sapiens
US-10-258-898A-5437

Query Match 82.9%; Score 34; DB 6; Length 310;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLP 5
|||||
DB 23 GGWLP 27

RESULT 6
PCT-US03-19871-10
Sequence 10, Application PC/TUS0319871
GENERAL INFORMATION:
APPLICANT: Eli Lilly and Company
TITLE OF INVENTION: Novel Proteins and Their Uses
FILE REFERENCE: X-15815
CURRENT APPLICATION NUMBER: PCT/US03/19871
CURRENT FILING DATE: 2003-07-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 10
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)-(312)
OTHER INFORMATION: LP395
PCT-US03-19871-10

Query Match 82.9%; Score 34; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLP 5
|||||
DB 272 GGWLP 276

RESULT 7
US-10-408-765A-1125
Sequence 1125, Application US/10408765A
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Wainock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1125
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-1125

Query Match 82.9%; Score 34; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLP 5
|||||
DB 7 GGWLP 11

RESULT 8
US-10-401-397A-2
Sequence 2, Application US/10401397A
GENERAL INFORMATION:
APPLICANT: Perl, Kriehna G.
APPLICANT: Moffett, Serge
APPLICANT: Adrian, Daniel
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION AND TREATMENT OF ELEVATED
TITLE OF INVENTION: INTRAOCULAR PRESSURE AND RELATED CONDITIONS
FILE REFERENCE: 4518/1M674US1
CURRENT APPLICATION NUMBER: US/10/401,397A
CURRENT FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: US 60/367,513
PRIOR FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-401-397A-2

Query Match 82.9%; Score 34; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLP 5
|||||
DB 305 GGWLP 309

RESULT 9
US-10-292-798-1274
Sequence 1274, Application US/10292798
GENERAL INFORMATION:
APPLICANT: SUGA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1274
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1274

Query Match 82.9%; Score 34; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLP 5
|||||
DB 305 GGWLP 309

RESULT 10
PCT-US03-19871-8
; Sequence 8, Application PC/TUS0319871
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: Novel Proteins and Their Uses
; FILE REFERENCE: X-15815
; CURRENT APPLICATION NUMBER: PCT/US03/19871
; CURRENT FILING DATE: 2003-07-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(328)
; OTHER INFORMATION: LP394
PCT-US03-19871-8

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
|||||
Db 288 GGMLP 292

RESULT 11
PCT-US03-19871-6
; Sequence 6, Application PC/TUS0319871
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: Novel Proteins and Their Uses
; FILE REFERENCE: X-15815
; CURRENT APPLICATION NUMBER: PCT/US03/19871
; CURRENT FILING DATE: 2003-07-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(371)
; OTHER INFORMATION: LP393
PCT-US03-19871-6

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
|||||
Db 288 GGMLP 292

RESULT 12
US-10-603-114-5455
; Sequence 5455, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
5

; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5455
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-5455

Query Match
Best Local Similarity 82.9%; Score 34; DB 6; Length 416;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
|||||
Db 175 GGMLP 179

RESULT 13
PCT-US03-19871-2
; Sequence 2, Application PC/TUS0319871
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: Novel Proteins and Their Uses
; FILE REFERENCE: X-15815
; CURRENT APPLICATION NUMBER: PCT/US03/19871
; CURRENT FILING DATE: 2003-07-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(486)
; OTHER INFORMATION: LP391
PCT-US03-19871-2

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
|||||
Db 272 GGMLP 276

RESULT 14
US-10-617-573-12
; Sequence 12, Application US/10617573
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul L.
; APPLICANT: Grimaldi, J.Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Li, Hanzhong
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Hymowitz, Sarah
; APPLICANT: Tumas, Daniel
; APPLICANT: Starovasnik, Melissa.
; APPLICANT: Vanlookeren, Menno
; APPLICANT: Vandlen, Richard
; APPLICANT: Watanabe, Colin
; APPLICANT: Williams, P.Mickey
; APPLICANT: Wood, William
; APPLICANT: Yansura, Daniel
; TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES THEREOF
; FILE REFERENCE: P1381RJC14(US)
; CURRENT APPLICATION NUMBER: US/10/617,573
; CURRENT FILING DATE: 2003-07-11

PRIOR APPLICATION NUMBER: US/10/000.157
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/138387
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/172096
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 60/175481
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 60/191007
PRIOR FILING DATE: 2000-03-21
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 12
LENGTH: 502
TYPE: PRT
ORGANISM: Homo Sapien
US-10-617-573-12

Query Match 82.9%; Score 34; DB 6; Length 502;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGWLP 5
Db 288 GGWLP 292

RESULT 15
PCT-US03-19871-4
Sequence 4, Application PC/TUS0319871
GENERAL INFORMATION:
APPLICANT: Eli Lilly and Company
TITLE OF INVENTION: Novel Proteins and Their Uses
FILE REFERENCE: X-15815
CURRENT APPLICATION NUMBER: PCT/US03/19871
CURRENT FILING DATE: 2003-07-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 531
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(531)
OTHER INFORMATION: LP392
PCT-US03-19871-4

Query Match 82.9%; Score 34; DB 1; Length 531;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGWLP 5
Db 317 GGWLP 321

Search completed: August 20, 2003, 12:45:15
Job time : 2.73494 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 4.19277 Seconds
(without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GGWLPY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : PTR_76:*

1: ptr1:*\n2: ptr2:*\n3: ptr3:*\n4: ptr4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	178	2	ORF12 - Agrobacter
2	39	95.1	167	2	hypothetical prote
3	39	95.1	172	2	hypothetical integ
4	39	95.1	235	2	conserved hypotet
5	39	95.1	1057	2	outer capsid prote
6	37	90.2	352	2	NADH dehydrogenase
7	37	90.2	637	2	T26593
8	37	90.2	647	2	T43952
9	37	90.2	759	2	T44142
10	36	87.8	596	2	T52331
11	35	85.4	150	2	S177062
12	35	85.4	251	2	T10467
13	35	85.4	254	2	T15032
14	35	85.4	263	2	T35032
15	35	85.4	308	2	D70875
16	35	85.4	342	2	AG2713
17	35	85.4	352	2	D97495
18	35	85.4	379	2	D81515
19	35	85.4	379	2	H81741
20	35	85.4	379	2	A71479
21	35	85.4	406	2	F84152
22	35	85.4	415	2	A86599
23	35	85.4	415	2	C72026
24	35	85.4	422	2	T40003
25	35	85.4	517	2	T37225
26	35	85.4	536	2	G88954
27	35	85.4	1329	2	D87226
28	35	85.4	1330	2	B70836
29	35	85.4	1928	2	JS0610

30	35	85.4	2831	2	T31419	cyclic beta 1-2 g1
31	35	85.4	2867	2	AG3481	cellobiose-phospho
32	34	82.9	106	2	S26461	Ig heavy chain V r
33	34	82.9	140	2	S46351	vpx protein - siml
34	34	82.9	167	2	A97760	NADH2 dehydrogenas
35	34	82.9	244	1	T01091	NADH2 dehydrogenas
36	34	82.9	246	2	F84412	oxidoeductase [lm
37	34	82.9	247	2	T06543	MADS box protein -
38	34	82.9	253	2	C83148	conserved hypotet
39	34	82.9	287	2	B70829	probable unaA2 pro
40	34	82.9	295	2	H87216	mycolic acid synth
41	34	82.9	322	2	A13395	NADH2 dehydrogenas
42	34	82.9	325	1	DNMTU1	NADH2 dehydrogenas
43	34	82.9	325	1	S49576	NADH2 dehydrogenas
44	34	82.9	325	1	B40358	NADH2 dehydrogenas
45	34	82.9	326	2	S59101	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

I39722

ORF12 - Agrobacterium rhizogenes plasmid pR18196

C:Species: Agrobacterium rhizogenes

C>Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 08-Oct-1999

C:Accession: I39722

R:Hansen, G.; Larrieu, M.; Vaubert, D.; Tempe, J.; Biermann, B.J.; Montoya, A.L.; Chl Proc. Natl. Acad. Sci. U.S.A. 88, 7763-7767, 1991

A:Title: Agrobacterium rhizogenes pR18196 T-DNA: Mapping and DNA sequence of function

A:Reference number: I39720; MUID:91352070; PMID:1909028

A:Accession: I39722

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-178 <RES>

A:Cross-references: GB:M60490; NID:g142245; PIDN:AAA22096.1; PID:g142248

C:Genetics:

A:Genome: Plasmid

A>Note: encoded within the T-DNA (transferred DNA) segment of the plasmid; this segme

t disease

Query Match

Best Local Similarity 100.0%; Score 41; DB 2; Length 178;

Matches 6; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGWLPY 6

76 GGWLPY 81

RESULT 2

hypothetical protein NMB1733 [imported] - Neisseria meningitidis (strain MC58 serogro

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: D81049

R:Retzlaff, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Qiu, H.; Yamathavan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizzo, M.

Science 287, 1809-1815, 2000

A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: AB1000; MUID:20175755; PMID:10710307

A:Accession: D81049

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-167 <RET>

A:Cross-references: GB:AE002523; GB:AE002098; NID:g7226980; PIDN:AAF42078.1; PID:g722

C:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1733

C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

Query Match 95.1%; Score 39; DB 2; Length 167;
Best Local Similarity 83.3%; Pred. No. 7.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMWLPY 6
DB 136 GGMWMPY 141

RESULT 3

hypoetical integral membrane protein NMA1989 [imported] - Neisseria meningitidis (stra
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: B81828
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
N:Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: B81828
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:97380371; PIDN:CAB85209.1; PID:9738062
A:Experimental source: serogroup A, strain z2491
C:Genetics:

A:Gene: NMA1989
C:Superfamily: Neisseria meningitidis hypothetical protein NMA1733

Query Match 95.1%; Score 39; DB 2; Length 172;
Best Local Similarity 83.3%; Pred. No. 7.4;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMWLPY 6
DB 141 GGMWMPY 146

RESULT 4

B83012
conserved hypothetical protein PA5071 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: B83012
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: B83012
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-235 <STO>
A:Cross-references: GB:AE004920; GB:AE004091; NID:9951358; PIDN:AAG08456.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:

A:Gene: PA5071

Query Match 95.1%; Score 39; DB 2; Length 235;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMWLPY 6
DB 196 GGMWIPY 201

RESULT 5

S47151
outer capsid protein VP2 - African horse sickness virus (serotype 3)
C:Species: African horse sickness virus

A:Variety: serotype 3
C:Date: 23-Nov-1994 #sequence_revision 03-Nov-1995 #text_change 20-Sep-1999
C:Accession: S47151
R:Vrede, F.T.; Huismans, H.

submitted to the EMBL Data Library, September 1993
A:Description: The cloning, characterization and expression of the gene that encodes
A:Reference number: S47151
A:Accession: S47151

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1057 <VR>
A:Cross-references: EMBL:Z26316; NID:9499374; PIDN:CAAB1225.1; PID:9499375
A:Experimental source: serotype 3
C:Genetics:

A:Map position: segment 2
C:Superfamily: African horse sickness virus VP2 protein

Query Match 95.1%; Score 39; DB 2; Length 1057;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMWLPY 6
DB 891 GGMWIPY 896

RESULT 6

E90175
NADH dehydrogenase subunit H (Nuoh) Nuoh [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: E90175
R:Sheng, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Ch
Jong, J.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90175

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <KUR>
A:Cross-references: GB:AE006641; NID:913813470; PIDN:AAK40660.1; GSPDB:GN00155
C:Genetics:

A:Gene: Nuoh
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1

Query Match 90.2%; Score 37; DB 2; Length 352;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGMWLPY 6
DB 284 GGMWLPY 289

RESULT 7

T26593
hypothetical protein Y32F6A.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26593
R:Barlow, K.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20241
A:Accession: T26593

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-637 <WIL>

A:Cross-references: EMBL:AL021474; PIDN:CAA16307.1; GSPDB:GN00023; CESP:Y32F6A.2
A:Experimental source: clone Y32F6A
C:Genetics:
A:Gene: CESP:Y32F6A.2

A:Map position: 5
A:introns: 45/3; 275/3; 376/2; 453/3; 500/1; 606/3

Query Match
Best Local Similarity 90.2%; Score 37; DB 2; Length 637;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGMLPY 6
|||||
DB 111 GGMPY 116

RESULT 8

T43952
hypothetical protein DR2 [imported] - human herpesvirus 6
C:Species: human herpesvirus 6

A:Variety: strain HST
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43952; T44059
R:Seigawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawe

J. Virol. 73, 8053-8063, 1999
A:Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and

A:Reference number: 222732; MUID:99412319; PMID:10482554
C:Accession: T43952

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-647 <I>S>
A:Cross-references: EMBL:AB021506; NID:g4995977; PIDN:BAAT8213.1; PID:g4995980

A:Experimental source: strain HST; pop. variant B
A:Accession: T44059

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-647 <I>S>
A:Cross-references: EMBL:AB021506; NID:g4995977; PIDN:BAAT8213.1; PID:g4995980

A:Experimental source: strain HST; pop. variant B
C:Genetics:

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

A:Gene: DR2, DR2R

Query Match
Best Local Similarity 90.2%; Score 37; DB 2; Length 647;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLPY 6
|||||
DB 235 GGMLPF 240

RESULT 9

T44142
DRI protein [imported] - human herpesvirus 6 (strain Z29)
C:Species: human herpesvirus 6

A:Variety: strain Z29
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T44142; T44239
R:Domínguez, G.; Dambaugh, T.R.; Stamey, F.R.; Demhurst, S.; Inoue, N.; Pellett, P.E.

J. Virol. 73, 8040-8052, 1999
A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with human

A:Reference number: 222734; MUID:99412318; PMID:10482553
C:Accession: T44142

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-759 <DOM>
A:Cross-references: EMBL:AF157706; PIDN:AA049614.1

A:Experimental source: strain Z29; variant B
A:Genetics: GN1

A:Accession: T44239
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-759 <DO2>

A:Cross-references: EMBL:AF157706; PIDN:AA049682.1
A:Experimental source: strain Z29; variant B

A:Genetics: GN2

C:Genetics: <GN1>

A:Gene: DRI
A:Map position: 583
A:introns: 87/1
A:Genetics: <GN2>
A:Gene: DRI
A:Map position: 153904
A:introns: 87/1

Query Match
Best Local Similarity 90.2%; Score 37; DB 2; Length 759;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLPY 6
|||||
DB 348 GGMLPF 353

RESULT 10

T52331
pectinesterase (EC 3.1.1.11) [imported] - Salix gilgiana
N:Alternate names: pectin methylsterase

C:Species: Salix gilgiana
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000

C:Accession: T52331
R:Futamura, N.; Mori, H.; Kouchi, H.; Shinohara, K.

Plant Cell Physiol. 41, 16-26, 2000
A:Title: Male flower-specific expression of genes for polygalacturonase, pectin methy

A:Reference number: 226040
A:Accession: T52331

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-596 <PUT>
A:Cross-references: EMBL:AB029461; PIDN:BA089480.1

A:Genetics: SGPME1
C:Superfamily: pectinesterase

C:Keywords: carboxylic ester hydrolase

Query Match
Best Local Similarity 87.8%; Score 36; DB 2; Length 596;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLPY 6
|||||
DB 526 GGMLPW 531

RESULT 11

S17702
hypothetical protein 1 - Azospirillum brasilense
C:Species: Azospirillum brasilense

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
C:Accession: S17702

R:Zimmer, W.; Aparicio, C.; Elmerich, C.
Mol. Gen. Genet. 229, 41-51, 1991

A:Title: Relationship between tryptophan biosynthesis and indole-3-acetic acid produc
A:Reference number: S17702; MUID:91375449; PMID:1896020

C:Accession: S17702
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-150 <I>M>

A:Cross-references: EMBL:X57853; NID:g48840; PIDN:CAA40983.1; PID:g48841

Query Match
Best Local Similarity 85.4%; Score 35; DB 2; Length 150;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGMLPY 6
|||||
DB 40 GGMSPY 45

RESULT 12

T00656
MADS box protein AGL9 - Arabidopsis thaliana
N:Alternate names: protein F316.19
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 29-Oct-1999
C:Accession: T00656
R:Feederstein, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo, I.; Vysotskaya, V.S.; Yu, G.; Becker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, February 1998
A:Reference number: 214197
A:Accession: T00656
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-251 <FE0>
A:Cross-references: EMBL:AC002396; NID:92749918; PIDN:AAC00586.1; PID:92829878; GSPDB:GN
C:Genetics:
A:Gene: ATSP:F316.19; AGL9
A:Map position: 1
A:introns: 62/2; 90/3; 113/2; 146/3; 160/3; 174/3; 222/2
C:Superfamily: transcription factor sqwa; serum response factor DNA-binding domain homol
C:Keywords: DNA binding; nucleus; transcription regulation
F:2-57/Domain: serum response factor DNA-binding domain homology <SRF>
Query Match 85.4%; Score 35; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GMLPY 6
DB 242 GMLPY 246

RESULT 13
T10467
MADS box protein D - white mustard
N:Alternate names: MADS-box protein AGL9 homolog
C:Species: Sinapis alba (white mustard)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10467
R:Bonhomme, F.; Sommer, H.; Bernier, G.; Jacquemard, A.
Plant Mol. Biol. 34, 573-582, 1997
A>Title: Characterization of SAMADS D from Sinapis alba suggests a dual function of the
A:Reference number: 217032; M0ID:97390682; PMID:9247539
A:Accession: T10467
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-254 <BON>
A:Cross-references: EMBL:Y08626; NID:g1617210; PIDN:CAA69916.1; PID:g1617211
C:Superfamily: transcription factor sqwa; serum response factor DNA-binding domain homol
F:2-57/Domain: serum response factor DNA-binding domain homology <SRF>
Query Match 85.4%; Score 35; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GMLPY 6
DB 245 GMLPY 249

RESULT 14
T35032
probable hydroxylase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000
C:Accession: T35032
R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: 221565
A:Accession: T35032
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-263 <SEE>

A:Cross-references: EMBL:AL079355; PIDN:CA845588.1; GSPDB:GN00070; SCORDB:SC4C6.24C
A:Experimental source: strain A3(2)
A:Genetics:
A:Gene: SCORDB:SC4C6.24C
C:Superfamily: Mycobacterium hypothetical protein Rv0911
Query Match 85.4%; Score 35; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GMLPY 6
DB 205 GMLPY 209

RESULT 15
D70875
probable PE protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70875
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; M0ID:98295987; PMID:9634230
A:Accession: D70875
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-308 <COL>
A:Cross-references: GB:AL010186; GB:AL123456; NID:93261493; PIDN:CA15849.1; PID:e120
A:Experimental source: strain H37Rv
A:Genetics:
A:Gene: PE

Query Match 85.4%; Score 35; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GMLPY 6
DB 223 GMLPY 227

Search completed: August 20, 2003, 12:42:26
Job time : 6.19277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 2.27711 Seconds
(Without alignments) 123.912 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GGLWLPY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwIsbProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	39	95.1	1057 1 VP2_AHSV3	089508 african hor
2	35	85.4	150 1 YTRL_AZOBH	P26943 azospirillum
3	35	85.4	251 1 SEB3_ARATH	022456 arabidopsis
4	35	85.4	254 1 AGS3_SINAI	004067 sinapis alb
5	35	85.4	506 1 CP7B_HUMAN	075881 homo sapien
6	35	85.4	507 1 CP7B_MOUSE	060991 mus musculu
7	35	85.4	1051 1 VP2_AHSV6	071024 african hor
8	35	85.4	1928 1 LPH_RAT	002401 rattus norv
9	34	82.9	167 1 NUOE_RICCN	092109 rickettsia
10	34	82.9	244 1 NUHM_WHEAT	022769 arabidopsis
11	34	82.9	325 1 NUHM_WHEAT	001148 triticum ae
12	34	82.9	326 1 NUHM_WHEAT	P48898 chondrus cr
13	34	82.9	328 1 NUHM_WHEAT	P26845 marchantia
14	34	82.9	331 1 NUHM_WHEAT	P31839 oenothera b
15	34	82.9	401 1 NUHM_WHEAT	094762 drosophila
16	34	82.9	434 1 NUHM_WHEAT	033180 mycobacteri
17	34	82.9	499 1 NUHM_WHEAT	093163 mus musculu
18	34	82.9	502 1 NUHM_WHEAT	094762 drosophila
19	34	82.9	622 1 KUP_ECOLI	094762 drosophila
20	33	80.5	166 1 LED2_THEMA	094762 drosophila
21	33	80.5	208 1 LED2_THEMA	094762 drosophila
22	33	80.5	331 1 YB94_METTH	094762 drosophila
23	33	80.5	354 1 YB94_METTH	094762 drosophila
24	33	80.5	395 1 YB94_METTH	094762 drosophila
25	33	80.5	501 1 YB94_METTH	094762 drosophila
26	33	80.5	1060 1 VP2_AHSV4	P51542 oryctolagus
27	32	78.0	135 1 Y40A_TREPA	P32553 african hor
28	32	78.0	309 1 Y40A_TREPA	P58007 terebrantia p
29	32	78.0	372 1 Y40A_TREPA	P54524 bacillus su
30	32	78.0	441 1 Y40A_TREPA	000155 lactococcus
31	32	78.0	446 1 Y40A_TREPA	094762 drosophila
32	32	78.0	486 1 Y40A_TREPA	P43581 saccharomyc
33	32	78.0	610 1 ARL1_YEAST	P25628 saccharomyc

34	32	78.0	1047 1 DHE2_NEUCR	P00365 neurospora
35	31	75.6	360 1 GP25_HUMAN	O00155 homo sapien
36	31	75.6	378 1 O33A_DROME	P81914 drosophila
37	31	75.6	408 1 PGK_PYRAE	O82W66 pyrobaculum
38	31	75.6	416 1 ENO_METTH	O26149 methanobact
39	31	75.6	489 1 C128_MYCTU	O59512 mycobacteri
40	31	75.6	495 1 AVNA_ASPPA	O12732 aspergillus
41	31	75.6	519 1 LNT_SYNY3	P74055 synechocyst
42	31	75.6	522 1 NUSM_MYTED	O00232 mytilus edu
43	31	75.6	559 1 YH01_ECOLI	P37659 escherichia
44	31	75.6	663 1 DUS8_ECOLI	O09112 mus musculu
45	31	75.6	663 1 TKT1_ECOLI	P27302 escherichia

ALIGNMENTS

RESULT 1	VP2_AHSV3	STANDARD;	PRT; 1057 AA.
AC	089508;		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DE	Outer capsid protein VP2.		
GN	S2 OR L2.		
OS	African horse sickness virus 3 (AHSV-3) (African horse sickness virus		
OS	(serotype 3)).		
OC	Viruses; dsRNA viruses; Reoviridae; Orbivirus.		
OX	NCBI_TaxID=117204;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95088612; PubMed=7996157;		
RA	Wrede F.T., Huisman H.;		
RT	"Cloning, characterization and expression of the gene that encodes		
RT	the major neutralization-specific antigen of African horsesickness		
RT	virus serotype 3."		
RL	J. Gen. Virol. 75:3629-3633(1994).		
CC	- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)		
CC	WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE		
CC	MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.		
CC	- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; U01832; AAA62184.1; -		
DR	EMBL; Z26316; CA81225.1; -		
DR	PIR; S47151; S47151.		
DR	InterPro; IPR001742; Orbl_VP2.		
DR	Pfam; PF00898; Orbl_VP2; 1.		
KW	Coat protein.		
SEQUENCE	1057 AA; 123063 MW; 2880D3DB091C06C4 CRC64;		
Query Match	95.1%; Score 39; DB 1; Length 1057;		
Best Local Similarity	83.3%; Pred. No. 24;		
Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
QY	1 GGLWLPY 6		
DB	891 GGLWLPY 896		
RESULT 2	YTRL_AZOBH	STANDARD;	PRT; 150 AA.
ID	YTRL_AZOBH		
AC	P26943;		
DT	01-AUG-1992 (Rel. 23, Created)		

```

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 16.3 Kda protein in TRP5 5' region (ONR1).
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Azospirillum.
OX NCBI_TaxID=192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP7;
RX MEDLINE=91375449; PubMed=1896020;
RA Zimmer W., Aparicio C., Elmerich C.;
RT "Relationship between tryptophan biosynthesis and indole-3-acetic
RT acid production in Azospirillum: Identification and sequencing of a
RT trpGDC cluster."
RL Mol. Gen. Genet. 229:41-51(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL; X57853; CAA40983.1;
DR PIR; S17702; S17702.
KW Hypothetical protein.
SQ SEQUENCE 150 AA: 16323 MW: DLE473075B44B0B2 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 150;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 40 GGWSPY 45

RESULT 3
SEPI_ARATH STANDARD; PRT; 251 AA.
ID SEPI_ARATH STANDARD; PRT; 251 AA.
AC 022456; 086W04;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Developmental protein SEPALLATA3 (Agamous-like MADS box protein AGL9).
CN SEPI OR AGL9 OR AT1G24260 OR F316.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RA Mandel M.A., Yanofsky M.F.;
RT "The Arabidopsis AGL9 MADS box gene is expressed in young flower
RT primordia."
RL Sex. Plant Reprod. 11:22-28(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altsch F.H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buell E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.U., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,

```

```

RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Matli R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakono H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tolium M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE OF 114-251 FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAPFs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC)."
RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX PubMed=9418042;
RA Fan H.-Y., Hu Y., Tudor M., Ma H.;
RT "Specific interactions between the K domains of AG and AGLs, members
RT of the MADS domain family of DNA binding proteins."
RL Plant J. 12:999-1010(1997).
RN [5]
RP CHARACTERIZATION.
RX PubMed=10821278;
RA Pelaz S., Ditta G.S., Baumann E., Wisman E., Yanofsky M.F.;
RT "B and C floral organ identity functions require SEPALLATA MADS-box
RT genes."
RL Nature 405:200-203(2000).
RN [6]
RP CHARACTERIZATION.
RX PubMed=11206550;
RA Honma T., Goto K.;
RT "Complexes of MADS-box proteins are sufficient to convert leaves into
RT floral organs."
RL Nature 409:525-529(2001).
CC -1- FUNCTION: Probable transcription factor active in inflorescence
CC development and floral organogenesis. Functions with
CC SEPALLATA1/AGL2 and SEPALLATA2/AGL4 to ensure proper development
CC of petals, stamens and carpels and to prevent the indeterminate
CC growth of the flower meristem. Interacts with APERALA1, AGAMOUS or
CC APERALA3/PISILLATA to form complexes, that could be involved in
CC genes regulation during floral meristem development.
CC -1- SUBUNIT: Heterodimer with API or AG capable of binding to CARG-box
CC sequences. Binds AP3/PI to form a ternary complex.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- DEVELOPMENTAL STAGE: Expressed early during flower development
CC within petals, stamens, and carpels
CC -1- MISCELLANEOUS: Triple mutations in the SEPI, SEPI2 and SEPI3 genes
CC result in the replacement of the stamens and petals by sepals and
CC of the carpels by a new mutant flower with sepaloid organs.
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS.
CC -1- SIMILARITY: Contains 1 K-box dimerization domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL; AF015552; AAB67832.1;
DR EMBL; AC002396; AAC00586.1;
DR EMBL; AK118696; BRC43290.1; ALT_INIT.
DR PIR; T00656; T00656.
DR HSSP; P11746; TMN6.
DR TRANSFAC; T03032;
DR InterPro; IPR002487; TF_Kbox.
DR InterPro; IPR002100; TF_MADSbox.

```


DR Pfam; PF01486; K-box; 1.
 DR Pfam; PF00319; SRF-TF; 1.
 DR PRINTS; PR00404; MADSDOMAIN.
 DR SMART; SM00432; MADS; 1.
 DR PROSITE; PS00350; MADS_BOX_1; 1.
 DR PROSITE; PS50066; MADS_BOX_2; 1.
 KW Flowering; transcription regulation; Activator; Developmental protein;
 KM Nuclear protein; DNA-binding; Coiled coil.
 FT DOMAIN 3 57 MADS.
 FT DOMAIN 100 172 K-BOX.
 FT DOMAIN 94 177 COILED COIL (POTENTIAL).
 SQ SEQUENCE 251 AA; 29066 MW; 0057CABD5F1AFC40 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 251;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 DB 242 GWLPY 246

RESULT 4
 AGL9-STNAL STANDARD; PRT; 254 AA.
 AC AGL9-STNAL
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Agamous-like MADS box protein AGL9 homolog (MADS D).
 GN AGL9.
 OS Sinapis alba (white mustard) (Brassica hirta).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Sinapis.
 NC NCBI_TaxID=3728;
 RX MEDLINE=97390682; PubMed=9247539;
 RA Bonhomme F., Sommer H., Bernier G., Jacquard A.;
 RT "Characterization of SAMDS D from Sinapis alba suggests a dual
 RT function of the gene: in inflorescence development and floral
 RT organogenesis.";
 RL Plant Mol. Biol. 34:573-573(1997).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ACTIVE IN INFLORESCENCE
 CC DEVELOPMENT AND FLORAL ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
 CC FACTORS.
 CC -1- SIMILARITY: Contains 1 K-box dimerization domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y08626; CAA69916.1; -
 DR PIR; T10467; T10467.
 DR HSSP; P11746; 1MM.
 DR TRANSFAC; T03170; -
 DR InterPro; IPR002487; TF_Rbox.
 DR InterPro; IPR002100; TF_MADSbox.
 DR Pfam; PF01486; K-box; 1.
 DR Pfam; PF00319; SRF-TF; 1.
 DR PRINTS; PR00404; MADSDOMAIN.
 DR SMART; SM00432; MADS; 1.
 DR PROSITE; PS00350; MADS_BOX_1; 1.
 DR PROSITE; PS50066; MADS_BOX_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein.

FT DOMAIN 3 57 MADS.
 FT DOMAIN 100 172 K-BOX.
 SQ SEQUENCE 254 AA; 29603 MW; EDD65155060BFD46 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 254;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLPY 6
 DB 245 GWLPY 249

RESULT 5
 CYP7B_HUMAN STANDARD; PRT; 506 AA.
 ID CYP7B_HUMAN
 AC 075881; Q9UNE5;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 7B1 (oxygensterol 7-alpha-hydroxylase) (EC 1.14.13.-).
 GN CYP7B1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBI_TaxID=9606;
 RX MEDLINE=99021668; PubMed=9802883;
 RA Setchell K.D.R., Schwarz M., O'Connell N.C., Lund E.G., Davis D.L.,
 RA Lathie R., Thompson H.R., Tyson W.R., Sokol R.J., Russell D.W.;
 RT "Identification of a new inborn error in bile acid synthesis: mutation
 RT of the oxygensterol 7-alpha-hydroxylase gene causes severe neonatal liver
 RT disease.";
 RL J. Clin. Invest. 102:1690-1703(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RX MEDLINE=20057832; PubMed=10588945;
 RA Wu Z.L., Martin K.O., Javitt N.B., Chiang J.Y.L.;
 RT "Structure and functions of human oxygensterol 7alpha-hydroxylase cDNAs
 RT and gene CYP7B1.";
 RL J. Lipid Res. 40:2195-2203(1999).
 CC -1- PATHWAY: Conversion of cholesterol to bile acids; first (rate-
 CC limiting) step.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: Brain, testis, ovary, prostate, liver, colon,
 CC kidney, and small intestine.
 CC DISEASE: Defects in CYP7B1 are a cause of neonatal giant cell
 CC hepatitis (MIM:231100). Patients present severe cholestasis,
 CC cirrhosis and liver synthetic failure.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF029403; AAC95426.1; -
 DR EMBL; AF127090; AAD20021.1; -
 DR EMBL; AF176805; AAK11850.1; -
 DR EMBL; AF176800; AAK11850.1; JOINED.
 DR EMBL; AF176801; AAK11850.1; JOINED.
 DR EMBL; AF176802; AAK11850.1; JOINED.
 DR EMBL; AF176803; AAK11850.1; JOINED.
 DR EMBL; AF176804; AAK11850.1; JOINED.
 DR Genew; HGNC:2652; CYP7B1.
 DR MIM; 603711; -
 DR MIM; 231100; -

DR GO: GO:0008396; F:oxysterol 7-alpha-hydroxylase activity; TAS.
 DR GO: GO:0006699; P:bile acid biosynthesis; TAS.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; FALSE_NEG.
 DR Oxidoreductase: Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Cholesterol metabolism.
 FT METAL 449 449 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 324 324 R->H (IN REF. 1).
 SQ SEQUENCE 506 AA; 58255 MW; 07D3D4B801B6DBD9 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 506;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
 DB 51 GWLPY 55

RESULT 6
 CP7B_MOUSE STANDARD; PRT: 507 AA.
 ID CP7B_MOUSE
 AC Q60991;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 7B1 (Oxysterol 7-alpha-hydroxylase) (EC 1.14.13.-)
 DE (Hc7-1).
 GN CYP7B1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Liver;
 RX MEDLINE=96102021; PubMed=8530364;
 RA Stapleton G., Steel M., Richardson M., Mason J.O., Rose K.A.,
 RA Morris R.G., Lath R.;
 RT "A novel cytochrome P450 expressed primarily in brain.";
 RL J. Biol. Chem. 270:29739-29745(1995).
 CC -1- PATHWAY: Conversion of cholesterol to bile acids; first (rate-limiting) step.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN; ALSO EXPRESSED IN LIVER AND KIDNEY.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: U36993; AAA92615.1; -
 DR MGD: MGI:104978; CYP7B1.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; FALSE_NEG.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Cholesterol metabolism.
 FT METAL 447 447 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 507 AA; 56417 MW; AB30502F935B7ED6 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 507;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
 DB 49 GWLPY 53

RESULT 7
 VP2_AHSV6 STANDARD; PRT: 1051 AA.
 ID VP2_AHSV6
 AC 071024;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Outer capsid protein VP2.
 GN 52 OR 12.
 OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus (serotype 6)).
 OS Viruses; dsRNA viruses; Reoviridae; Orbivirus.
 OX NCBI_TaxID=86060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98278331; PubMed=9617769;
 RA Williams C.F., Inoue T., Lucas A.M., Zanotto P., Roy P.;
 RT "The complete sequence of four major structural proteins of African horse sickness virus serotype 6: evolutionary relationships within and between the Orbiviruses.";
 RL Virus Res. 53:53-73(1998).
 CC -1- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5) WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
 CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: AF021235; AAC40994.1; -
 DR InterPro: IPR001742; Orbl_VP2.
 DR Pfam: PF00898; Orbl_VP2; 1.
 KW Coat protein.
 SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9E389F4B5F CRC64;

Query Match 85.4%; Score 35; DB 1; Length 1051;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GWLPY 6
 DB 886 GWLPY 890

RESULT 8
 LPH_RAT STANDARD; PRT: 1928 AA.
 ID LPH_RAT
 AC Q02401; Q63712; Q63719;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 30-JUL-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lactase-phenolizin hydrolase precursor (lactase-glycosylceramidase) (includes: Lactase (EC 3.2.1.108); Phenolizin hydrolase (EC 3.2.1.62)).
 GN LCT OR LPH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Intestine;
 RX MEDLINE=91365258; PubMed=1909681;
 RA Dulac I., Boukamel R., Mantel N., Semenza G., Raul F., Freund J.-N.;

RT "Sequence of the precursor of intestinal lactase-phenolizin hydrolase
from fetal rat.";
RL Gene 103:275-276(1991).
RN [2]
RP SEQUENCE OF 1-192 FROM N.A.
RC STRAIN-Sprague-Dawley;
RX MEDLINE-93091244; PubMed-1339333;
RA Boukamel R., Freund J.-N.;
RT "The rat LPH gene 5' region: comparative structure with the human
gene.";
RL DNA Seq. 3:119-121(1992).
CC -1- FUNCTION: LPH SPLITS LACTOSE IN THE SMALL INTESTINE.
CC -1- CATALYTIC ACTIVITY: Lactose + H(2)O = D-glucose + D-galactose.
CC -1- CATALYTIC ACTIVITY: Glycosyl-N-acetylphingosine + H(2)O = a sugar +
N-acetylphingosine.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. BRUSH BORDER.
CC -1- TISSUE SPECIFICITY: INTESTINE.
CC -1- DOMAIN: THE SEQUENCE EXHIBITS 4 REGIONS (I-IV) OF INTERNAL
HOMOLOGY; THEREFORE LPH MIGHT HAVE EVOLVED BY TWO CYCLES OF
PARTIAL GENE DUPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
or send an email to license@slb.ch).

DR EMBL: X56748; CAA40070.1; -
DR EMBL: X56747; CAA40069.1; -
DR EMBL: L04635; AAA41539.1; -
DR PIR: J50610; J50610.
DR HSSP: P26205; 1C8G.
DR InterPro: IPR001360; Glyco_hydro_1.
DR Pfam: PF00232; Glyco_hydro_1; 5.
DR PRINTS: PR00151; GLHYDRLASE.
DR ProDom: PD000650; Glyco_hydro_1; 4.
DR PROSITE: PS00572; GLYCOSYL_HYDROL_F1_1; 1.
DR PROSITE: PS00653; GLYCOSYL_HYDROL_F1_2; 2.
DR KW Hydrolase; Glycosidase; Zymogen; Signal; Transmembrane; Repeat.
FT SIGNAL 1
FT PROPEP 20
FT CHAIN 868
FT DOMAIN 22
FT TRANSMEM 1884
FT DOMAIN 1903
FT DOMAIN 1928
FT DOMAIN 89
FT REPEAT 89
FT REPEAT 365
FT REPEAT 886
FT REPEAT 1371
FT ACT_SITE 1067
FT ACT_SITE 1274
FT ACT_SITE 1539
FT ACT_SITE 1750
FT CONFLICT 7
FT CONFLICT 113
FT CONFLICT 207
SQ SEQUENCE 1928 AA; 217266 MW; 56DDCAAC4ACAE85 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 1928;
Best Local Similarity 100.0%; Pred. No. 2,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GWLP 6
DB 1089 GWLP 1093
RESULT 9
NUOE_RICCN

ID NUOE_RICCN STANDARD; PRT; 167 AA.
AC 0921D9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-quinone oxidoreductase chain E (EC 1.6.99.5) (NADH dehydrogenase
I, chain E) (NDH-1, chain E).
GN NUOE OR RC0481.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiase; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Malish 7;
RX MEDLINE-21442074; PubMed-11557893;
RA Ogata H., Audic S., Rensato-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissensbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
electrons transferred, four hydrogen ions are translocated across
the cytoplasmic membrane), and thus conserves the redox energy in
a proton gradient (By similarity).
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- CORRECTOR: Binds 1 2Fe-2S cluster (Potential).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 24 kDa SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
or send an email to license@slb.ch).

DR EMBL: AE008611; AAL03019.1; -
DR PIR: A97760; A97760.
DR InterPro: IPR002023; Cmplx1_24kDa.
DR Pfam: PF01257; complex1_24kDa; 1.
DR ProDom: PD003859; Cmplx1_24kDa; 1.
DR PROSITE: PS01099; COMPLEX1_24K; 1.
DR Oxidoreductase; NAD: Quinone; Metal-binding; Iron-sulfur; Iron;
KW 2Fe-2S; Complete proteome.
FT METAL 91
FT METAL 96
FT METAL 132
FT METAL 136
SQ SEQUENCE 167 AA; 19156 MW; DE33C1F29451EB4 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWLP 5
DB 46 GWLP 50
RESULT 10
NUHM_ARATH
ID NUHM_ARATH STANDARD; PRT; 244 AA.
AC 022769;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase 24 kDa subunit, mitochondrial precursor
DE (EC 1.6.5.3) (EC 1.6.99.3).
GN At4G02580 OR T10P11.14.
OS Arabidopsis thaliana (Mouse-ear cress).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_TaxID=3702;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=cv. Columbia;
 CC MEDLINE=20083488; PubMed=10617198;
 CC MAYER K.F.X., SCHUELLER C., WAMBUTT R., MURPHY G., VOLCKAERT G.,
 CC POHL T., DUESSTHOEF A., STIEKEMA W., ENLICH K.-D., TERRY N.,
 CC HARRIS B., ANSORGE W., BRANDT P., GRIVELL L.A., RIEGER M.,
 CC WEICHELGAERTNER M., DE SIMONE V., OBERMAIER B., MACHE R., MUELLER M.,
 CC KREIS M., DEISENY M., PUIGDOMENECH P., WATSON M., SCHMIDTKEIN T.,
 CC REICHERT B., PORTETELLE D., PEREZ-ALONSO M., BOUTRY M., BANCROFT I.,
 CC VOS P., HOEHNSEL J., ZIMMERMANN W., WEDLER H., RIDLEY P.,
 CC LANGHAM S.-A., MCCULLAGH B., BILHAM L., ROBEN J.,
 CC VAN DER SCHUREN J., GYMONPREZ B., CHANG Y.-J., VANDENBUSCH F.,
 CC BREKEN M., WELTJENS I., VOET M., BASTIAENS I., AERT R., DEFEOE E.,
 CC WEITZENEGGER T., BOEHE G., RAMSPERGER U., HILBERT H., BRAUN M.,
 CC HOLZER E., BRANDT A., PETERS S., VAN STAVEREN M., DIKSE W.,
 CC MOOLMAN P., KLEIN LANKHORST R., ROSE M., HAUF J., KETTER P.,
 CC BENEISER S., HEMPEL S., FELDPAUSCH M., LAMBERTH S., VAN DEN DAELE H.,
 CC DE KEYSER A., BYSSHAERT C., GIELEN J., VILLARROEL R., DE CIERQ R.,
 CC VAN MONTAGU M., ROGERS J., CRONIN A., QUAIL N., BRAY-ALLEN S.,
 CC CLARK L., DOGGETT J., HALL S., KAY M., LENNARD N., MCILAY K., MAYES R.,
 CC PETTEIT A., RAJADREAM M.A., LYNE M., BENES V., RECHMANN S.,
 CC BORKOVA D., BLOECKER H., SCHARFE M., GRIMM M., LOENHART T.-H.,
 CC DOSE S., DE HAAN M., MAARSE A.C., SCHAEFER M., MUELLER-AUER S.,
 CC NEUMANN S., ARGITLON A., VITALE D., LIQUORI R., PLAVANDI E.,
 CC MASSENET O., QUIGLEY F., CLABAUD G., MUENDLEIN A., FELBER R.,
 CC SCHNABL S., HILLER R., SCHMIDT W., LECHARNY A., AUBOURG S.,
 CC CHEDOR F., COOKE R., BERGER C., MONFORT A., CASACUBERTA E.,
 CC GIBBONS T., WEBER N., VANDENBOL M., BARGUES M., TEROL J., TORRES A.,
 CC PEREZ-PEREZ A., PUNELLE B., BENT E., JOHNSON S., TACON D., JESSE T.,
 CC HEIJNEN L., SCHWARTZ S., SCHOLLER P., HEBER S., FRANCES P., BIELKE C.,
 CC REISHMAN D., HAASE D., LEMCKE K., MEWES H.-W., STOKER S.,
 CC ZACCARIEL P., BEVAN M., WILSON R.K., DE LA BASTIDE M., HABERMAN K.,
 CC PANELL L., DEHLA N., GNOJ L., SCHULTZ K., HUANG E., SPIEGEL L.,
 CC SEKHON M., MURRAY J., SHEET P., CORDES M., ABU-THREIDEN J.,
 CC STONEKING T., KALICKI J., GRAVES T., HARMON G., EDWARDS J.,
 CC LATREILLE P., COURTNEY L., CLOUD J., ABBOTT A., SCOTT K., JOHNSON D.,
 CC MINX P., BENTLEY D., FULLON B., MILLER N., GRECO T., KEMP K.,
 CC KRAMER J., FULLON L., MARDIS E., DANTE M., PEPIN K., HILLER L.,
 CC NELSON J., SPIETH J., RYAN E., ANDREWS S., GEISEL C., LAYMAN D.,
 CC DU H., ALI J., BERHOFF A., JONES K., DRONE K., COTTON M., JOSHU C.,
 CC ANTONIOU B., ZIDANIC M., STRONG C., SUN H., LAMAR B., YORDAN C.,
 CC MA P., ZHONG J., PRESTON R., VIL D., SHEKHER M., MATERO A., SHAN R.,
 CC SWADY I.K., O'SHAUGHNESSY A., RODRIGUEZ M., HOFFMAN J., TILL S.,
 CC GRANAT S., SHOHY N., HASEGAWA A., HAMEED A., LODHI M., JOHNSON A.,
 CC CHEN E., MARA M., MARLENGEN R., MCCOMBIE W.R.;
 CC "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 CC thaliana."
 CC Nature 402:769-777(1999).
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE. COMPONENT OF THE FLAVOPROTEIN-SULFUR (FP)
 CC FRAGMENT OF THE ENZYME (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -1- COFACTOR: Binds 1 2Fe-2S cluster (potential).
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
 CC THIS IS A COMPONENT OF THE FLAVOPROTEIN FRACTION.
 CC -1- SUBCELLULAR LOCATION: MATRIX AND CYTOPLASMIC SIDE OF THE
 CC MITOCHONDRIAL INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 24 kDa SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; AC002330; AAC78260.1; -
 CC EMBL; AL161494; CAB80751.1; -
 CC PIR; T01091; T01091.
 CC HSSP; 066511; 1F37.
 CC DR InterPro: IPR002023; Cmp1x1_24kDa.
 CC Pfam; PF01257; complex_24kd. 1.
 CC PRODOM; PD003859; Cmp1x1_24kDa; 1.
 CC PROSITE; PS01099; COMPLEX1_24K; 1.
 CC OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion; Transist peptide;
 CC Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
 CC TRANSIT 1
 CC CHAIN ? 244 MITOCHONDRION (POTENTIAL).
 CC NADH-UBIQUINONE OXIDOREDUCTASE 24 kDa
 CC SUBUNIT.
 CC METAL 119 119 IRON-SULFUR (2FE-2S) (POTENTIAL).
 CC METAL 124 124 IRON-SULFUR (2FE-2S) (POTENTIAL).
 CC METAL 160 160 IRON-SULFUR (2FE-2S) (POTENTIAL).
 CC METAL 164 164 IRON-SULFUR (2FE-2S) (POTENTIAL).
 CC SEQUENCE 244 AA; 27182 MW; 27C95BF5884B12AC CRC64;
 CC Query Match 82.9%; Score 34; DB 1; Length 244;
 CC Best Local Similarity 100.0%; Pred. No. 47;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC 1 GSWLP 5
 CC 74 GSWLP 78
 CC -----
 CC RESULT 11
 CC NUIW.WHEAT STANDARD; PRT; 325 AA.
 CC ID NUIW.WHEAT
 CC AC 001148;
 CC DT 01-APR-1993 (Rel. 25, Created)
 CC DT 01-APR-1993 (Rel. 25, Last sequence update)
 CC DT 15-JUN-1999 (Rel. 38, Last annotation update)
 CC DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 CC GN NDI OR NAD1.
 CC OS Triticum aestivum (Wheat).
 CC OC Mitochondrion.
 CC CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 CC CC Triticeae; Triticum.
 CC CC NCBI_TaxID=4565;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Thatcher;
 CC RX MEDLINE=91208683; PubMed=1902143;
 CC RA Chappelaine Y., Bonen L.;
 CC RT "The wheat mitochondrial gene for subunit I of the NADH dehydrogenase
 CC complex: a trans-splicing model for this gene-in-pieces.";
 CC Cell 65:465-472(1991).
 CC RL Cell 65:465-472(1991).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
 CC -1- CAUTION: Positions 1, 72, 103, 146, 164, 167, 179, 191, 193, 203,
 CC 225, 245 and 248 are modified by RNA editing.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; X57968; CAA41034.1; ALT_SEQ.
 CC EMBL; X57967; CAA41034.1; JOINED.
 CC EMBL; X57966; CAA41034.1; JOINED.
 CC EMBL; X57965; CAA41034.1; JOINED.
 CC PIR; A38489; DNMWTU1.
 CC InterPro: IPR001694; Resp_NADH_dhl.
 CC Pfam; PF00146; NADHdh; 1.

DR PROSITE: PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE: PS00668; COMPLEX1_ND1_2; 1.
 KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane;
 KM RNA editing.
 SQ SEQUENCE 325 AA; 35932 MW; BBD2BC96498CF53 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 325;
 Best Local Similarity 100.0%; Pred. NO. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
 |||||
 DB 249 GGMLP 253

RESULT 12
 ID NUIM_CHOCR STANDARD; PRT; 326 AA.

AC P48898:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN NDI OR NAD1.
 OS Chondrus crispus (Carrageen).
 OC Mitochondrion.
 OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartineae;
 OC Chondrus.
 CC NCBI_TaxID=2769;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Apices;
 RX MEDLINE=95341681; PubMed=7616569;
 RA Leblanc C., Boyen C., Richard O., Grienenberger J.M.,
 RA Kleareg B.;
 RT "Complete sequence of the mitochondrial DNA of the rhodophyte
 RT Chondrus crispus (Gigartinales). Gene content and genome
 RT organization";
 RL J. Mol. Biol. 250:484-495(1995).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL; 247547; CA87617.1; -;
 DR PIR; S59101; S59101.
 DR InterPro; IPR001694; Resp_NADH_dhl.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 326 AA; 36447 MW; 6D33B81B09173EE2 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. NO. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
 |||||
 DB 250 GGMLP 254

RESULT 13
 ID NUIM_MARPO STANDARD; PRT; 328 AA.
 AC P26845:
 DT 01-AUG-1992 (Rel. 23, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN NDI OR NAD1.
 OS Marchantia polymorpha (Liverwort).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
 OC Marchantiopsida; Marchantiales; Marchantaceae;
 OC Marchantia.
 CC NCBI_TaxID=3197;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=92114051; PubMed=1731062;
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
 RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohama K.;
 RT "Gene organization deduced from the complete sequence of liverwort
 RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
 RT mitochondrial genome.";
 RL J. Mol. Biol. 223:1-7(1992).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL; M68929; AAC09438.1; -;
 DR PIR; S25993; S25993.
 DR InterPro; IPR001694; Resp_NADH_dhl.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KM Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 328 AA; 36518 MW; F0F6F009567264DB CRC64;

Query Match 82.9%; Score 34; DB 1; Length 328;
 Best Local Similarity 100.0%; Pred. NO. 61;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMLP 5
 |||||
 DB 252 GGMLP 256

RESULT 14
 ID NUIM_OENBE STANDARD; PRT; 331 AA.
 AC P31839:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN NDI OR NAD1.
 OS Oenothera lutea (Evening primrose).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Myrtales; Onagraceae; Oenothera.
 CC NCBI_TaxID=3950;
 RN [1]
 RP SEQUENCE FROM N.A. AND RNA EDITING.

RX MEDLINE=91208684; PubMed=1850322;
 RA Wisinger B., Schuster W., Brennicke A.;
 RT "Trans splicing in Oenothera mitochondria: nadl mRNAs are edited in
 RT exon and trans-splicing group II intron sequences.";
 RL Cell 65:473-482(1991).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-7 IS THE INITIATOR.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63033; -; NOT_ANNOTATED_CDS.
CC Interpro; IPR001694; Resp_NADH_ch1.
CC Pfam; PF00146; NADHdh; 1.
CC PROSITE; PS00667; COMPLEXI_ND1_1; 1.
CC PROSITE; PS00668; COMPLEXI_ND1_2; 1.
CC OXidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane;
CC KW editing.
CC SEQUENCE 331 AA; 36667 MW; 26764C0FF001A82A CRC64;

Query Match 82.9%; Score 34; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSWLP 5
Db 255 GSWLP 259

RESULT 15
088A_DROME STANDARD; PRT; 401 AA.
ID 088A_DROME
AC 09VENZ;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative odorant receptor 88a.
GN OR88A OR CG14360.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Caylor S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reijntj K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Syntas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN 12
RP CONCEPTUAL TRANSLATION.
RA Robertson H.M.;
RT Unpublished observations (May-2001).
CC -1- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT
CC RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED
CC RECEPTORS.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003703; AAF55018.1; ALT_INIT.
CC FLYBASE; FBgn0038203; Or88a.
CC Interpro; IPR004117; 7tm_6.
CC Pfam; PF02949; 7tm_6; 1.
CC DR Hypothetical protein; Transmembrane; G-protein coupled receptor;
CC KW Olfaction; Multigene family.
CC FT DOMAIN 1 26
CC FT TRANSMEM 27 47
CC FT DOMAIN 48 52
CC FT TRANSMEM 53 73
CC FT DOMAIN 74 142
CC FT TRANSMEM 143 163
CC FT DOMAIN 164 191
CC FT TRANSMEM 192 212
CC FT DOMAIN 213 277
CC FT TRANSMEM 278 298
CC FT DOMAIN 299 303
CC FT TRANSMEM 304 324
CC FT DOMAIN 325 370
CC FT TRANSMEM 371 391
CC FT DOMAIN 392 401
CC SEQUENCE 401 AA; 47088 MW; 8B55679940963623 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSWLP 5
Db 179 GSWLP 183

Search completed: August 20, 2003, 12:35:01
Job time : 4.38822 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21; Search time 11.3494 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-33

Perfect score: 41

Sequence: 1 GGWLPY 6

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL.23:*
1: sp_Archea:*
2: sp_Bacteria:*
3: sp_Fungi:*
4: sp_Human:*
5: sp_Invertebrate:*
6: sp_Mammal:*
7: sp_Mhc:*
8: sp_Organelle:*
9: sp_Phage:*
10: sp_Plant:*
11: sp_Rodent:*
12: sp_Virus:*
13: sp_Vertebrate:*
14: sp_Unclassified:*
15: sp_Virus:*
16: sp_Bacteria:*
17: sp_Archea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	178	2	044194
2	39	95.1	167	16	09JY55
3	39	95.1	172	16	09JY37
4	39	95.1	235	16	09H082
5	39	95.1	1057	12	088B84
6	37	90.2	352	17	0980H2
7	37	90.2	397	16	08Y326
8	37	90.2	620	12	089893
9	37	90.2	637	5	045915
10	37	90.2	647	12	09W964
11	37	90.2	715	10	08VXB6
12	37	90.2	759	12	09PMX1
13	36	87.8	216	10	094DV0
14	36	87.8	310	16	08X0H8
15	36	87.8	382	2	08KN98
16	36	87.8	575	16	08CJL7

17	36	87.8	596	10	09MB66	09mb6 salix gligl
18	35	85.4	102	10	08GM04	08gm04 arabidopsis
19	35	85.4	169	12	08V3H6	08v3h6 swinepox v1
20	35	85.4	235	16	08Z0Z0	08z0z0 salmonella
21	35	85.4	250	10	08L9R5	08l9r5 arabidopsis
22	35	85.4	263	16	09XAL6	09xal6 streptomyces
23	35	85.4	289	12	09DXG8	09dxg8 beak and fe
24	35	85.4	289	12	09DXF9	09dx98 beak and fe
25	35	85.4	289	12	09DXF3	09dx98 beak and fe
26	35	85.4	289	12	09DXG5	09dx98 beak and fe
27	35	85.4	289	12	09YUD3	09yud3 beak and fe
28	35	85.4	289	12	09DXF6	09dx98 beak and fe
29	35	85.4	289	12	09DXH0	09dx98 beak and fe
30	35	85.4	289	12	09DXH2	09dx98 beak and fe
31	35	85.4	293	12	09DXG2	09dx98 beak and fe
32	35	85.4	293	12	09LEK3	09lek3 goose circo
33	35	85.4	293	12	08BCC3	08bcc3 goose circo
34	35	85.4	293	12	08AYY2	08ayy2 goose circo
35	35	85.4	293	12	08AYV0	08ayv0 goose circo
36	35	85.4	293	12	08AYV9	08ayv9 goose circo
37	35	85.4	299	12	090235	090235 beak and fe
38	35	85.4	308	16	050428	050428 mycobacteri
39	35	85.4	320	16	08RE46	08re46 fusobacteri
40	35	85.4	332	16	092MT1	092mt1 rhizobium m
41	35	85.4	340	17	08TL25	08tl25 methanocarc
42	35	85.4	352	16	08UGC6	08ugc6 agrobacteri
43	35	85.4	360	2	08GAC2	08gac2 lyngbya ma
44	35	85.4	379	16	09PLK0	09plk0 chlamydia m
45	35	85.4	379	16	09KIT9	09kit9 chlamydia p

ALIGNMENTS

RESULT 1
ID 044194 PRELIMINARY; PRT; 178 AA.
AC 044194;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ORF12.
OS Agrobacterium rhizogenes.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91352070; PubMed=1809028;
RA Hansen G., Larride M., Vaubert D., Tempe J., Biermann B.J.,
RA Montoya A.L., Chilton M.-D., Brevet J.,
RT "Agrobacterium rhizogenes PR18196 T-DNA: Mapping and DNA sequence of
RT functions involved in mannopine synthesis and hairy root
RT differentiation."
RT functions involved in mannopine synthesis and hairy root
RL Proc. Natl. Acad. Sci. U.S.A. 88:7763-7767(1991).
RL EMBL; M60490; AAA22096.1; -
DR Interpro: IPR006064; Glycosylase.
DR Interpro: IPR006065; Glyco_hydro.41.
DR Pfam: PF02027; K01B_R01C; 1.
DR PRINTS: PR00746; GLHYDRASE41.
SQ SEQUENCE 178 AA; 22A160ED64C45A8 CRC64;

Query Match 100.0%; Score 41; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 76 GGWLPY 81

RESULT 2
09JY55

ID 09JY55 PRELIMINARY; PRT; 167 AA.
 AC 09JY55;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical protein NMB1733.
 GN NMB1733.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
 Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
 Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58".
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002523; AAF42078.1; -
 DR TIGR: NMB1733; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 167 AA; 19854 MW; 173631A183372CDE CRC64;

Query Match 95.1%; Score 39; DB 16; Length 167;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
 |||:|
 DB 136 GGWMPY 141

RESULT 3
 O9JY37 PRELIMINARY; PRT; 172 AA.
 AC 09JY37;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical integral membrane protein.
 GN NMA1989.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parhill J., Achman M., James K.D., Bentley S.D., Churcher C.,
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 Jagers K., Leach S., Moule S., Mungall K., Quail M.A.,
 Rajendram M.A., Rutherford K.M., Simmonds M., Skelton J.,
 Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis 22491".
 RL Nature 404:502-506(2000).
 DR EMBL: AL162757; CAB85209.1; -
 KW Complete proteome.
 SQ SEQUENCE 172 AA; 20428 MW; 0363DA56B0692A1 CRC64;

Query Match 95.1%; Score 39; DB 16; Length 172;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
 |||:|
 DB 141 GGWMPY 146

RESULT 4
 O9HUB2 PRELIMINARY; PRT; 235 AA.
 AC 09HUB2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein PA5071.
 GN PA5071.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 Barber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen".
 RL Nature 406:959-964(2000).
 DR EMBL: AE004920; AAC08456.1; -
 DR InterPro: IPR004382; Cons_hypoth46.
 DR TIGRFRMS: TIGR00046; TIGR00046; 2.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 235 AA; 26097 MW; E874C787AAC00DF CRC64;

Query Match 95.1%; Score 39; DB 16; Length 235;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
 |||:|
 DB 196 GGWMPY 201

RESULT 5
 O8B8V4 PRELIMINARY; PRT; 1057 AA.
 AC 08B8V4;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Major outer capsid protein VP2.
 GN VP2.
 OS African horseshoe virus.
 OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
 OX NCBI_TaxID=40050;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Potieter A.C., Cloete M., van Dijk A.A.;
 RT "A first full outer capsid protein data set in the Orbiviridae:
 RT Cloning, sequencing, expression and analysis of a complete set of
 RT full-length outer capsid VP2-genes of the nine African horseshoe
 RT virus serotypes".
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY163330; AAN74570.1; -
 KW Complete proteome.
 SQ SEQUENCE 1057 AA; 123656 MW; 43EF1F41E37EE77C CRC64;

Query Match 95.1%; Score 39; DB 12; Length 1057;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMWLPY 6
 DB 891 GGMWLPY 896

RESULT 6

O980H2 PRELIMINARY: PRT; 352 AA.
 AC O980H2;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE NADH dehydrogenase subunit H (NUOH).
 GN NUOH OR SSO0325.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE-21332296; PubMed-11427726;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Welher C.C.Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: AE006667; AAK0660.1; -.
 DR InterPro: IPR001694; Resp_NADH_dhl.
 DR Pfam: PF00146; NADHdh; 1.
 KM Complete proteome.
 SQ SEQUENCE 352 AA; 39368 MW; F82B6D033396F828 CRC64;

Query Match Best Local Similarity 90.2%; Score 37; DB 17; Length 352;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMWLPY 6
 DB 284 GGMWLPY 289

RESULT 7

O8Y326 PRELIMINARY: PRT; 397 AA.
 AC O8Y326;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein RSC0155.
 GN RSC0155 OR RS01025.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM1100;
 RX MEDLINE-21681879; PubMed-11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Atilat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavié M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
 RL Nature 415:497-502(2002).
 DR EMBL: AL646057; CAD13683.1; -.

DR InterPro: IPR003788; DUF185.
 DR Pfam: PF02636; DUF185; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 397 AA; 42650 MW; 502B0D984A1477F5 CRC64;

Query Match Best Local Similarity 90.2%; Score 37; DB 16; Length 397;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGMWLPY 6
 DB 33 GGMWLPY 38

RESULT 8

O89893 PRELIMINARY: PRT; 620 AA.
 AC O89893;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical 67.1 kDa protein.
 GN DR2.
 OS Human herpesvirus 6.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 NCBI_TaxID=10368;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-9526321; PubMed-7747482;
 RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
 RA Martin M.E., Efsthaniou S., Craxton M., Macaulay H.A.,
 RT "The DNA sequence of human herpesvirus-6: structure, coding content,
 and genome evolution."
 RL Virology 209:29-51(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-90080132; PubMed-2152817;
 RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W.,
 RA Barrett B.G.,
 RT "Human herpesvirus 6 is closely related to human cytomegalovirus."
 RL J. Virol. 64:287-299(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-91237802; PubMed-1851860;
 RA Chang C.K., Balachandran N.,
 RT "Identification, characterization, and sequence analysis of a CDNA
 encoding a phosphoprotein of human herpesvirus 6."
 RL J. Virol. 65:2884-2894(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-9133007; PubMed-1651403;
 RA Teo I.A., Griffin B.E., Jones M.D.,
 RT "Characterization of the DNA polymerase gene of human herpesvirus 6."
 RL J. Virol. 65:4670-4680(1991).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-91226542; PubMed-1851522;
 RA Thomson B.J., Efsthaniou S., Honess R.W.,
 RT "Acquisition of the human adeno-associated virus type-2 rep gene by
 human herpesvirus type-6."
 RL Nature 351:78-80(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-91374590; PubMed-1654446;
 RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.,
 RT "Identification of a transactivating function mapping to the putative

RT immediate-early locus of human herpesvirus 6.";
 RL J. Virol. 65:5381-5390(1991).
 RN (17)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-92333249; PubMed-1321206;
 RA Efsthaliou S., Lawrence G.L., Brown C.M., Barrell B.G.;
 RT "Identification of homologues to the human cytomegalovirus US22 gene
 RL family in human herpesvirus 6.";
 RL J. Gen. Virol. 73:1661-1671(1992).
 RN (18)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-92148942; PubMed-1310766;
 RA Geng Y., Chandran B., Josephs S.F., Wood C.;
 RT "Identification and characterization of a human herpesvirus 6 gene
 RL segment that trans activates the human immunodeficiency virus type 1
 RT promoter.";
 RL J. Virol. 66:1564-1570(1992).
 RN (9)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93091236; PubMed-1333836;
 RA Gompels U.A., Cars A.L., Sun N., Arrand J.R.;
 RT "Infectivity determinants encoded in a conserved gene block of human
 RL herpesvirus-6.";
 RL DNA Seq. 3:25-39(1992).
 RN (10)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-92260671; PubMed-1374813;
 RA Neipel F., Ellinger K., Fleckenstein B.;
 RT "Gene for the major antigenic structural protein (p100) of human
 RL herpesvirus 6.";
 RL J. Virol. 66:3918-3924(1992).
 RN (11)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-92333248; PubMed-1321205;
 RA Thomson B.J., Honess R.W.;
 RT "The right end of the unique region of the genome of human herpesvirus
 RL 6 U1102 contains a candidate immediate early gene enhancer and a
 RL homologue of the human cytomegalovirus US22 gene family.";
 RL J. Gen. Virol. 73:1649-1660(1992).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93187613; PubMed-8383182;
 RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Fiume G.,
 RL Fleckenstein B.;
 RL "The glycoprotein B homologue of human herpesvirus 6.";
 RL J. Gen. Virol. 74:495-500(1993).
 RN (13)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93224882; PubMed-8385692;
 RA Gompels U.A., Carrigan D.R., Cars A.L., Arno J.;
 RT "Two groups of human herpesvirus 6 identified by sequence analyses of
 RL laboratory strains and variants from Hodgkin's lymphoma and bone marrow
 RL transplant patients.";
 RL J. Gen. Virol. 74:613-622(1993).
 RN (14)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93389439; PubMed-8397282;
 RA Liu D.X., Gompels U.A., Nicholas J., Belliot C.;
 RT "Identification and expression of the human herpesvirus 6 glycoprotein
 RL H and interaction with an accessory 40K glycoprotein.";
 RL J. Gen. Virol. 74:1847-1857(1993).
 RN (15)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-94025558; PubMed-7692666;

RA Liu D.X., Gompels U.A., Foa-Tomasi L., Campadelli-Fiume G.;
 RT "Human herpesvirus-6 glycoprotein H and L homologs are components of
 RL the gp100 complex and the gH external domain is the target for
 RL neutralizing monoclonal antibodies.";
 RL Virology 197:12-22(1993).
 RN (16)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93331710; PubMed-7687803;
 RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,
 RA Greenmoyer C., Damhaugh T.R.;
 RT "A strongly immunoreactive virion protein of human herpesvirus 6
 RL variant B strain 229: identification and characterization of the gene
 RL epitope.";
 RL Virology 195:521-531(1993).
 RN (17)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-93323202; PubMed-7687301;
 RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwattanpong S.,
 RA Chandran B.;
 RT "Identification and mapping of the gene encoding the glycoprotein
 RL complex gp82-gp105 of human herpesvirus 6 and mapping of the
 RL neutralizing epitope recognized by monoclonal antibodies.";
 RL J. Virol. 67:4611-4620(1993).
 RN (18)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-95146989; PubMed-7844567;
 RA Gompels U.A., Macaulay H.A.;
 RT "Characterization of human telomeric repeat sequences from human
 RL herpesvirus 6 and relationship to replication.";
 RL J. Gen. Virol. 76:451-458(1995).
 RN (19)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-94047392; PubMed-8230490;
 RA Dewhurst S., Dollard S.C., Pellett P.E., Damhaugh T.R.;
 RT "Identification of a lytic-phase origin of DNA replication in human
 RL herpesvirus 6B strain 229.";
 RL J. Virol. 67:7680-7683(1993).
 RN (20)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RA Nicholas J.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 RN (21)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-94118404; PubMed-8289364;
 RA Nicholas J., Martin M.;
 RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
 RL genome of human herpesvirus 6 encoding human cytomegalovirus immediate-
 RL early gene homologs and transactivating functions.";
 RL J. Virol. 68:597-610(1994).
 RN (22)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-94202284; PubMed-8151768;
 RA Schiwe U., Neipel F., Schreiner D., Fleckenstein B.;
 RT "Structure and transcription of an immediate-early region in the human
 RL herpesvirus 6 genome.";
 RL J. Virol. 68:2978-2985(1994).
 RN (23)
 RP SEQUENCE FROM N.A.
 RC STRAIN-U1102;
 RX MEDLINE-94181269; PubMed-8134119;
 RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
 RA Frenkel N., Rosenthal L.J.;

Query Match 90.28; Score 37; DB 12; Length 620;
 Best Local Similarity 83.38; Pred. No. 2.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
|||||

Db 217 GGWLPY 222

RESULT 9

O45915 PRELIMINARY; PRT; 637 AA.

AC O45915;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
GN Y32FE6A.2 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Felodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL021474; CNA16307.1;
DR WormRep: Y32FE6A.2; CE16609.
DR InterPro: IPR000175; Na/ntra_n_sypmt.
DR Pfam: PF00209; SNF; 1.
DR ProDom: PD000448; Na/ntra_n_sypmt; 1.
DR ProSITE: PS50267; NA_NEUTROTAN_SYP_3; 1.
SQ SEQUENCE 637 AA; 72425 MW; 9C317888C0AE013 CRC64;

Query Match Best Local Similarity 90.2%; Score 37; DB 5; Length 637;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGWLPY 6
|||||

Db 111 GGWLPY 116

RESULT 10

O9W9G4 PRELIMINARY; PRT; 647 AA.

AC O9W9G4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN DR2R.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HST;
RA Isegawa Y., Mukai T., Nakano K., Kagawa M., Chen J., Mori Y.,
RA Sunagawa T., Sashihara J., Zou P., Kosuge H., Yamagishi K.;
RT "A comparison of the complete DNA sequences between human herpesvirus-
6 variant A and B."
RL J. Virol. 0:0-0(1999).
DR EMBL: AB021506; BAA78320.1;
DR EMBL: AB021506; BAA78323.1;
DR InterPro: IPR002965; P-rich_extensn.
DR InterPro: IPR003360; US22.

DR Pfam: PF02393; US22; 1.
DR PRINTS: PR01217; PRICEXTENSN.
SQ SEQUENCE 647 AA; 69466 MW; 21845BA4DAD1568 CRC64;

Query Match Best Local Similarity 90.2%; Score 37; DB 12; Length 647;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
|||||

Db 235 GGWLPY 240

RESULT 11

O8VXB6 PRELIMINARY; PRT; 715 AA.

AC O8VXB6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
GN Putative potassium transporter.
OS HAK4.
OC Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Nipponbare;
RA Banuelos M.A., Garciladebas B., Rodriguez-Navarro A.;
RT "Inventory of HAK transporters in rice."
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ227975; CAD20996.1;
DR Gramene: O8VXB6;
DR InterPro: IPR003855; K+-transporter.
DR Pfam: PF02705; K_trans; 1.
DR TIGRFAMs: TIGR00794; ktp; 1.
SQ SEQUENCE 715 AA; 79662 MW; B708C0A6748D77EF CRC64;

Query Match Best Local Similarity 90.2%; Score 37; DB 10; Length 715;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGWLPY 6
|||||

Db 489 GGWLPY 494

RESULT 12

O9PWX1 PRELIMINARY; PRT; 759 AA.

AC O9PWX1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN DRL.
OS Human herpesvirus 6B.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=32604;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-429;
RA MEDLINE=99412318; PubMed=10482553;
RA Dominguez G., Danbaugh T.R., Stamey F.R., Dewhurst S., Inoue N.,
RA Pellett P.E.;
RT "Human herpesvirus 6B genome sequence: coding content and comparison
with human herpesvirus 6A."
RL J. Virol. 73:8040-8052(1999).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-229;
RA Pellett P.E., Dominguez G., Dambaugh T.R., Stamey F.R., Dewhurst S.,
RA Inoue N.,
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF157706; AAD49682.1;
DR EMBL; AF157706; AAD49682.1;
DR InterPro: IPR002965; P_Rich_extensn.
DR InterPro: IPR003360; US22.
DR Pfam: PF02393; US22; 1.
DR PRINTS: PR01217; PRICHEXTENS.
SQ SEQUENCE 759 AA; 82564 MW; C3956651B1A28E28 CRC64;

Query Match
Best Local Similarity 90.2%; Score 37; DB 12; Length 759;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 348 GGWLPP 353

RESULT 13
O94DV0 PRELIMINARY; PRT; 216 AA.

AC O94DV0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE P0454H12.20 Protein.
GN P0454H12.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0454H12.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003255; BAB62587.1;
DR Gramene; O94DV0;
SQ SEQUENCE 216 AA; 23220 MW; 743AD9190D20A2FD CRC64;

Query Match
Best Local Similarity 87.8%; Score 36; DB 10; Length 216;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 153 GGWLPP 158

RESULT 14
O8XOH8 PRELIMINARY; PRT; 310 AA.

AC O8XOH8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable transcription regulator protein.
GN RSP1247 OR RS03191.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM11000;
RX MEDLINE-21681879; PubMed-11823852;

RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunne S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:487-502(2002)
DR EMBL; AL646083; CAD18398.1;
DR InterPro: IPR000005; HTMAAC.
DR Pfam; PF00165; HTH_ARAC; 2.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SMO0342; HTH_ARAC; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Plasmid, Complete Proteome
SQ SEQUENCE 310 AA; 33657 MW; 4A6D1F611686BEAD CRC64;

Query Match
Best Local Similarity 87.8%; Score 36; DB 16; Length 310;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 268 GGWLPP 273

RESULT 15
O8KN98 PRELIMINARY; PRT; 382 AA.

AC O8KN98;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE ORF 15.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-22053227; PubMed-12057956;
RA Raymond C.K., Sims E.H., Kas A., Spencer D.H., Kutayav N.V.,
RA Ivey R.G., Zhou Y., Kaul R., Clendenning J.B., Olson M.V.;
RT "Genetic Variation at the O-Antigen Biosynthetic Locus in Pseudomonas
aeruginosa.";
RL J. Bacteriol. 184:3614-3622(2002).
DR EMBL; AF498403; AAM27596.1;
SQ SEQUENCE 382 AA; 43581 MW; 565B3B2F274764D0 CRC64;

Query Match
Best Local Similarity 87.8%; Score 36; DB 2; Length 382;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGWLPY 6
DB 71 GGWLPP 76

Search completed: August 20, 2003, 12:40:35
Job time : 14.4605 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:12:21 ; Search time 13.1928 Seconds
(without alignments)
72.188 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIP 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 15876573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	24	AA156833	Apple LRPKml LRR s
2	34	100.0	108	AA153774	VL component of an
3	34	100.0	108	AAE08819	Human scfry lig ant
4	34	100.0	109	AA156836	Pinus radiata cell
5	34	100.0	144	AA156836	Human secreted pro
6	34	100.0	153	AA156836	Pinus radiata cell
7	34	100.0	155	AA156836	Eucaalyptus grandis
8	34	100.0	163	AA156836	Pinus radiata cell
9	34	100.0	224	AA156836	Pinus radiata cell

10	34	100.0	238	AA153775	An antibody with 1
11	34	100.0	307	AA156833	Protonibacterium
12	34	100.0	307	AA156833	Human novel protei
13	34	100.0	312	AA156833	Human novel protei
14	34	100.0	330	AA156833	Maize disease resi
15	34	100.0	334	AA156833	Drosophila melanog
16	34	100.0	705	AA156833	Pinus radiata cell
17	34	100.0	999	AA156833	Apple leucine-rich
18	34	97.1	94	AA156833	Human leucine-rich
19	34	94.1	295	AA156833	Human immune/haema
20	34	94.1	420	AA156833	Human novel protei
21	34	94.1	420	AA156833	Human novel protei
22	34	94.1	420	AA156833	Human novel protei
23	34	94.1	420	AA156833	Human novel protei
24	34	94.1	420	AA156833	Human novel protei
25	34	94.1	420	AA156833	Human novel protei
26	34	94.1	420	AA156833	Human novel protei
27	34	94.1	420	AA156833	Human novel protei
28	34	94.1	420	AA156833	Human novel protei
29	34	94.1	420	AA156833	Human novel protei
30	34	94.1	420	AA156833	Human novel protei
31	34	94.1	420	AA156833	Human novel protei
32	34	94.1	420	AA156833	Human novel protei
33	34	94.1	420	AA156833	Human novel protei
34	34	94.1	420	AA156833	Human novel protei
35	34	94.1	420	AA156833	Human novel protei
36	34	94.1	420	AA156833	Human novel protei
37	34	94.1	420	AA156833	Human novel protei
38	34	94.1	420	AA156833	Human novel protei
39	34	94.1	420	AA156833	Human novel protei
40	34	94.1	420	AA156833	Human novel protei
41	34	94.1	420	AA156833	Human novel protei
42	34	94.1	420	AA156833	Human novel protei
43	34	94.1	420	AA156833	Human novel protei
44	34	94.1	420	AA156833	Human novel protei
45	34	94.1	420	AA156833	Human novel protei

ALIGNMENTS

RESULT 1	AA156833	standard; peptide: 24 AA.
XX	AA156833;	
AC	AA156833;	
XX	11-APR-2000 (first entry)	
XX	Apple LRPKml LRR sequence #8.	
DE	Apple LRPKml LRR sequence #8.	
XX	LRPKml gene; leucine-rich protein kinase; apple; fungal pathogen;	
KW	Venturia Inaequalis; transgenic plant; plant defense; scab; LRR.	
XX		
OS	Malus domestica.	
XX		
PN	WO9964600-A1.	
XX		
PD	16-DEC-1999.	
XX		
PF	08-JUN-1999; 99WO-IT00165.	
XX		
PR	08-JUN-1998; 98IT-RM00367.	
XX		
PA	(AGRA-) IST AGRARIO DI SAN MICHELE ALL'ADIGE.	
XX		
PI	Cervone F, De Lorenzo G, Komjanc M;	
XX		
DR	WPI, 2000-147095/13.	
XX		
PT	New leucine-rich protein useful to increase plant resistance to fungal	
PT	pathogens, especially Venturia Inaequalis	
XX		

PS Disclosure; Fig 4; 45pp; English.

CC The invention relates to a LRPKml gene encoding a leucine-rich protein
 CC kinase from apple (Malus x domestica) cultivar Florida responsible for
 CC resistance to the fungal pathogen Venturia inaequalis. Vectors
 CC comprising the polynucleotide can be used to produce transgenic plants
 CC (especially apple species); resistant to fungal pathogens, especially V.
 CC inaequalis. For example, the gene sequence encoding the protein could be
 CC incorporated with the 35S promoter of the CamV cauliflower mosaic virus
 CC to introduce a plant defense mechanism against the pathogen, or the
 CC promoter sequence upstream from the gene could be used in chimeric
 CC constructs to promote sequences encoding for proteins which inhibit V.
 CC inaequalis. V. inaequalis is the most severe and economically important
 CC fungal pathogen of apples, causing a disease known as scab. Sequences
 CC AAY56826-848 represent apple LRPKml LRR sequences.

SO Sequence 24 AA;

Query Match 100.0%; Score 34; DB 21; Length 24;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 |||||
 Db 18 TGRIP 23

RESULT 2
 AAY53774
 ID AAY53774 standard; Peptide; 108 AA.

XX AAY53774;

DT 22-FEB-2000 (first entry)

DE VL component of an antibody with improved specificity for fibronectin.

XX scFv; antibody: ED-B domain epitope; fibronectin; marker;
 KW angiogenesis; vascular proliferation; diabetic retinopathy;
 KW age-related macular degeneration; tumour; immunosuppressive detection;
 KW blood coagulation; blood vessel occlusion; ocular angiogenesis;
 KW angiogenesis-related pathology.

XX Synthetic.

OS Homo sapiens.

PN WO958570-A2.

PD 18-NOV-1999.

PE 11-MAY-1999; 99WO-EP03210.

PR 11-MAY-1998; 98US-0075338.

PR 28-APR-1999; 99US-0300425.

PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.

PI Neri D, Tarli L, Viti F, Birchler M;

DR WPI; 2000-039074/03.

PT Fibronectin ED-B domain epitope specific antibodies and conjugate
 PT antibodies -

PS Claim 10; Page 38; 59pp; English.

CC The present sequence represents the VL component of a modified human
 CC scFv antibody which has specific affinity for a characteristic epitope
 CC of the ED-B domain of fibronectin. The affinity of the antibody for
 CC this epitope was improved by introducing a number of mutations in the
 CC complementarity determining region (CDR) residues located at the
 CC periphery of the binding site. The improved antibody is used for rapid
 CC targeting markers of angiogenesis, for detecting diseases characterized

CC by vascular proliferation, such as diabetic retinopathy, age-related
 CC macular degeneration or tumours. The antibody localizes the respective
 CC tissue within 3 to 4 hours after injection. It is used in
 CC immunosuppressive detection of angiogenesis and for diagnosis and
 CC therapy of tumours and diseases characterized by vascular proliferation.
 CC The antibody can be conjugated to a molecule which induces blood
 CC coagulation and blood vessel occlusion. These conjugates are used in
 CC the preparation of injectable compositions for the treatment of
 CC angiogenesis-related pathologies, especially caused by or associated
 CC with ocular angiogenesis.

SO Sequence 108 AA;

Query Match 100.0%; Score 34; DB 21; Length 108;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 |||||
 Db 92 TGRIP 97

RESULT 3
 AAE08819
 ID AAE08819 standard; Protein; 108 AA.

XX AAE08819;

DT 19-NOV-2001 (first entry)

DE Human scFv L19 antibody variable light chain region.

XX Human; scFv; single-chain variable antibody fragment; cancer; cytotoxic;
 KW coagulant; ED-B domain; fibronectin; tumour; ocular disorder; psoriasis;
 KW vascular proliferation; rheumatoid arthritis; blood vessel occlusion;
 KW angiogenesis; blood coagulation; variable light chain; VL.

XX Homo sapiens.

PN WO200162800-A1.

PD 30-AUG-2001.

PE 23-FEB-2001; 2001WO-EP02062.

PR 24-FEB-2000; 2000US-0512082.

PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.

PI Neri D, Tarli L, Viti F, Birchler M;

DR WPI; 2001-541701/60.

PT An antibody, with specific affinity for a characteristic epitope of the

PT ED-B domain of fibronectin for the treatment of diseases characterized

PT by vascular proliferation -

PS Claim 10; Page 36; 73pp; English.

CC The invention relates to an antibody with specific affinity for a
 CC characteristic epitope of the ED-B domain of fibronectin, where the
 CC antibody has improved affinity to ED-B. The invention also relates to
 CC conjugates comprising antibodies with a suitable photoactive molecule
 CC useful in the detection and/or coagulation of blood vessels. An antibody
 CC with improved affinity to the ED-B domain is useful for diagnosis and
 CC therapy of tumours and diseases characterized by vascular proliferation,
 CC cancer, rheumatoid arthritis, neo-vascularization associated ocular
 CC disorders and psoriasis. Treatment of angiogenesis related pathologies
 CC comprises the injection of conjugates comprising antibody and a molecule
 CC capable of inducing blood coagulation and blood vessel occlusion. The
 CC present sequence is single-chain variable antibody fragment (scFv) L19
 CC antibody variable light chain (VL) region related to the invention.

SQ Sequence 108 AA:

Query Match 100.0%; Score 34; DB 22; Length 108;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||||
 DB 92 TGRIPP 97

RESULT 4

AAB25366
 ID AAB25366 standard; Protein: 109 AA.

AC AAB25366;

DT 27-NOV-2000 (first entry)

DE Pinus radiata cell signalling involved protein SEQ ID NO:685.

XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism.

OS Pinus radiata.

XX WO200042171-A1.

PD 20-JUL-2000.

PE 11-JAN-2000; 2000WO-US00724.

PR 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strabala TJ, Muewenhulzen NJ;

XX WPI; 2000-476052/41.

PT Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 PT to external signals -

PS Claim 3; Page 314; 527pp; English.

XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.

SQ Sequence 109 AA:

Query Match 100.0%; Score 34; DB 21; Length 109;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||||
 DB 21 TGRIPP 26

RESULT 5

AAG00142
 ID AAG00142 standard; Protein: 144 AA.

AC AAG00142;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 4223.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PE 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

PR N-PSDB; AAC00148.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 13; SEQ ID 4223; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs with genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 144 AA:

Query Match 100.0%; Score 34; DB 21; Length 144;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||||
 DB 136 TGRIPP 141

RESULT 6

AAB25115
 ID AAB25115 standard; Protein: 153 AA.

AC AAB25115;

DT 27-NOV-2000 (first entry)

DE Pinus radiata cell signalling involved protein SEQ ID NO:83.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
XX plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
RW elongation; survival; disease resistance; nutrient metabolism.
XX
OS Pinus radiata.
XX
MO200042171-A1.
PN
PD 20-JUL-2000.
XX
PF 11-JAN-2000; 2000WO-US00724.
XX
PR 12-JAN-1999; 99US-0228986.
PR 01-NOV-1999; 99US-0162866.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strabala TJ, Nieuwenhuizen NJ;
XX
DR WPI: 2000-476052/41.
XX
PT Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals -
XX
PS Claim 3; Page 82-83; 527pp; English.
XX
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or
CC pine (*Pinus radiata*, also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
SQ Sequence 153 AA:

Query Match 100.0%; Score 34; DB 21; Length 153;
Best local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 TGRIPP 6
|||||
Db 126 TGRIPP 131

RESULT 7
AAB25326
ID AAB25326 standard; Protein; 155 AA.
XX
AC AAB25326;
XX
DT 27-NOV-2000 (first entry)
XX
DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:645.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;

```

KW elongation; survival; disease resistance; nutrient metabolism.
XX
XX Eucalyptus grandis.
OS
XX WO200042171-A1.
XX
PD 20-JUL-2000.
XX
XX 11-JAN-2000; 2000WO-US00724.
XX
XX 12-JAN-1999; 99US-0228986.
XX
XX 01-NOV-1999; 99US-0162866.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Strabala TJ, Nieuwenhuizen NJ;
XX
XX WPI: 2000-476052/41.
XX
XX Isolated polynucleotide encoding a polypeptide involved in cell
XX signaling used for generating transgenic plants with modified responses
XX to external signals -
XX
XX
XX Claim 3; Page 294; 527pp; English.
XX
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
XX and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
XX pine (Pinus radiata also known as Monterey pine). The protein sequences
XX are involved in cell signalling. The polynucleotide and protein
XX sequences can be used to modify the response of plant cells to external
XX signals e.g. environmental changes or pathogens during the growth and
XX development of a plant. They can be used to modify cell proliferation,
XX differentiation, elongation and survival, resistance to disease and
XX nutrient metabolism. Examples of modifications which can be produced are
XX altered fruit ripening and senescence of leaves and flowers e.g. to
XX delay senescence and prolong the life of cut flowers or enhance
XX senescence of reproductive organs to engineer sterile plants. Other
XX modifications can be used to delay senescence in selected cell types or
XX organs providing fruit and vegetables which have a longer shelf life
XX between harvest and consumption, or to decrease branching frequency in
XX forest tree species giving long stretches of valuable knot-free clear
XX wood which can be used in solid timber furniture and veneers.
XX
XX
XX Sequence 155 AA:
XX
XX Query Match 100.0%; Score 34; DB 21; Length 155;
XX Best Local Similarity 100.0%; Pred. No. 66;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TGRIPP 6
XX
XX 116 TGRIPP 121
XX
XX
XX RESULT 8
XX AAB25447
XX ID AAB25447 standard; Protein; 163 AA.
XX
XX AC AAB25447;
XX
XX DT 27-NOV-2000 (first entry)
XX
XX DE Pinus radiata cell signalling involved protein SEQ ID NO:766.
XX
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
XX plant cell signalling; modulation; transgenic plant; pathogen; growth;
XX environmental change; development; cell proliferation; differentiation;
XX elongation; survival; disease resistance; nutrient metabolism.
XX
XX Pinus radiata.
XX
XX OS
XX PN WO200042171-A1.
XX

```


PD 20-JUL-2000.
 XX
 PF 11-JAN-2000; 2000WO-US00724.
 XX
 PR 12-JAN-1999; 99US-0228986.
 PR 01-NOV-1999; 99US-0162866.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Strabala TJ, Nieuwenhuizen NT;
 XX
 DR WPI: 2000-476052/41.
 XX
 PT Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 PT to external signals -
 XX
 PS Claim 3; Page 355; 527pp; English.
 XX
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.
 CC
 XX Sequence 163 AA:
 XX
 Query Match 100.0%; Score 34; DB 21; Length 163;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TGRIPP 6
 DB 126 TGRIPP 131
 XX
 RESULT 9
 AAB25432
 ID AAB25432 standard; Protein; 224 AA.
 XX
 AC AAB25432;
 XX
 DT 27-NOV-2000 (first entry)
 XX
 DE Pinus radiata cell signalling involved protein SEQ ID NO:751.
 XX
 KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism.
 XX
 OS Pinus radiata.
 OS
 PN WO200042171-A1.
 PN
 PD 20-JUL-2000.
 PD
 PF 11-JAN-2000; 2000WO-US00724.
 PF
 PR 12-JAN-1999; 99US-0228986.
 PR 01-NOV-1999; 99US-0162866.
 PR

XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Strabala TJ, Nieuwenhuizen NT;
 XX
 DR WPI: 2000-476052/41.
 XX
 PT Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 PT to external signals -
 XX
 PS Claim 3; Page 347-348; 527pp; English.
 XX
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.
 CC
 XX Sequence 224 AA:
 XX
 Query Match 100.0%; Score 34; DB 21; Length 224;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TGRIPP 6
 DB 188 TGRIPP 193
 XX
 RESULT 10
 AAY53775
 ID AAY53775 standard; Protein; 238 AA.
 XX
 AC AAY53775;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE An antibody with improved specificity for fibronectin.
 XX
 KW scFv; antibody; ED-B domain epitope; fibronectin; marker;
 KW angiogenesis; vascular proliferation; diabetic retinopathy;
 KW age-related macular degeneration; tumour; immunosclerographic detection;
 KW blood coagulation; blood vessel occlusion; ocular angiogenesis;
 KW angiogenesis-related pathology.
 XX
 OS Synthetic.
 OS
 PN Homo sapiens.
 PN
 PD 18-NOV-1999.
 PD
 PF 11-MAY-1999; 99WO-EP03210.
 PF
 PR 11-MAY-1998; 98US-0075338.
 PR 28-APR-1999; 99US-0300425.
 PR
 PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
 PA
 PI Neel D, Taril L, Viti F, Birchler M;

XX WPI: 2000-039074/03.
XX Fibronectin ED-B domain epitope specific antibodies and conjugate
PT antibodies
PS
XX
PS
XX
CC Claim 10; Page -: 59pp; English.
CC The present sequence represents a modified human scfv antibody which
CC has specific affinity for a characteristic epitope of the ED-B domain
CC of fibronectin. The affinity of the antibody for this epitope was
CC improved by introducing a number of mutations in the complementarity
CC determining region (CDR) residues located at the periphery of the
CC binding site. The improved antibody is used for rapid targeting markers
CC of angiogenesis, for detecting diseases characterized by vascular
CC proliferation, such as diabetic retinopathy, age-related macular
CC degeneration or tumours. The antibody localizes the respective tissue
CC within 3 to 4 hours after injection. It is used in immunoscintigraphic
CC detection of angiogenesis and for diagnosis and therapy of tumours and
CC diseases characterized by vascular proliferation. The antibody can be
CC conjugated to a molecule which induces blood coagulation and blood
CC vessel occlusion. These conjugates are used in the preparation of
CC injectable compositions for the treatment of angiogenesis-related
CC pathologies, especially caused by or associated with ocular
CC angiogenesis.
CC Note: this sequence does not appear in the specification; it is an
CC amalgamation of the sequences given in claim 10.
CC
SQ Sequence 238 AA;
Query Match 100.0%; Score 34; DB 21; Length 238;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGRIPP 6
Db 222 TGRIPP 227
RESULT 11
AAU40893 standard; Protein; 307 AA.
XX
AC AAU40893;
XX
DT 13-FEB-2002 (first entry)
XX
DE Proionibacterium acnes immunogenic protein #1789.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Proionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI: 2001-616774/71.
DR N-PSDB; AAS59513.

XX Proionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
PS
XX
PS
XX
XX Example 1; SEQ ID No 2088; 1069pp; English.
CC Sequences AAU39105-AAU68017 represent Proionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 307 AA;
Query Match 100.0%; Score 34; DB 22; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGRIPP 6
Db 84 TGRIPP 89
RESULT 12
AAU14225 standard; Protein; 307 AA.
XX
AC AAU14225;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #96.
XX
KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytoskeletal; neuroprotective; vulnerary; noctropic;
KW anticonvulsant; antiepileptic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-451939/48.
DR N-PSDB; AAS22530.

XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 PS Example 4; Page 575; 894pp; English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

XX Sequence 307 AA;

Query Match 100.0%; Score 34; DB 22; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
 |||||
 Db 241 TGRIP 246

RESULT 13
 AAU14227

XX AAU14227 standard; Protein: 312 AA.

AC AAU14227;

XX 24-OCR-2001 (first entry)

XX Human novel protein #98.

XX Human: novel protein; Antianaemic; osteopathic; antiinflammatory;
 KM immunomodulatory; cytostatic; neuroprotective; vulnary; nocrotic;
 KM anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KM antibacterial; antiallergic; dermatological; haemostatic; antiastringent;
 KM thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KM Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KM tissue regeneration; immune disorder.

XX Homo sapiens.

XX WO200155437-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02623.

XX 25-JAN-2000; 2000US-0491404.

XX (HYSE-) HYSED INC.

PI Tang YT, Liu C, Dymnac RT;
 XX WPI: 2001-451939/48.
 DR N-PSDB; AAS22532.

XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 PS Example 4; Page 576-577; 894pp; English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

XX Sequence 312 AA;

Query Match 100.0%; Score 34; DB 22; Length 312;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
 |||||
 Db 246 TGRIP 251

RESULT 14

XX AAU00443 standard; Protein: 330 AA.

AC AAU00443;

XX 19-JUN-2001 (first entry)

XX Maize disease resistance enhancing protein ZMLR1-1.

XX Maize; plant disease resistance; crop; soybean; sunflower; sorghum;
 KM canola; wheat; alfalfa; cotton; rice; barley; millet; plant cell death;
 KM herbicide resistance; ZMLR1-1; Cf-2; Cf-9.

XX Zea mays.

XX WO200118061-A2.

XX 15-MAR-2001.

XX 06-SEP-2000; 2000WO-US24403.

XX 09-SEP-1999; 99US-0152988.

XX (PION-) PIONEER HI-BRED INT INC.

PI Simmons CR:

XX WPI: 2001-226742/23.
DR N-PSDB: AAS01013.

PT Novel isolated maize disease resistance polynucleotide useful for
PT increasing resistance in a plant to disease, controlling cell death,
PT and conferring resistance to herbicides -
PS Claim 11; Page 72; 90pp; English.

XX The present sequence representing maize ZmLRR1-1 protein is 1 of 7
CC novel disease resistance proteins (AAU00443-AAU00449). The ZmLRR1-1 is
CC a zea may gene homologue for a leucine-rich repeat (LRR) containing
CC disease resistance gene of the Cf-2 or Cf-9 type. These novel maize
CC disease resistance polynucleotides and polypeptides are useful for
CC enhancing disease resistance in crops and transgenic plants including
CC maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley or millet. The level of disease resistance protein is increased
CC by transforming a plant cell with a recombinant expression cassette
CC which comprises the disease resistance polynucleotide operably linked to
CC a promoter, or by culturing the plant cell under plant growing conditions
CC to produce a regenerated plant, or by inducing expression of the
CC polynucleotide to modulate the disease resistance protein in a plant.
CC The polynucleotides encoding the disease resistance proteins are useful
CC for increasing resistance in a plant to disease, controlling cell death,
CC and conferring resistance to herbicides. They are useful as probes or
CC amplification primers in the detection, quantification, or isolation
CC of gene transcripts. They can be used for recombinant expression of
CC their encoded polypeptides, as immunogens in the preparation and/or
CC screening of antibodies, and in sense or antisense suppression of the
CC polynucleotide in a host cell, tissue or plant.

SQ Sequence 330 AA:

Query Match 100.0%; Score 34; DB 22; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
DB 186 TGRIP 191

RESULT 15
ABB65742
ID ABB65742 standard; Protein; 334 AA.

XX ABB65742;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 24018.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.
DR

DR N-PSDB: ABL09845.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS Disclosure; SEQ ID NO 24018; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB157737-AB12072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 334 AA:

Query Match 100.0%; Score 34; DB 22; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
DB 231 TGRIP 236

Search completed: August 20, 2003, 12:33:49
Job time : 14.1928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:27:41 ; Search time 4.15663 Seconds
(without alignments)
61.075 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIP 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilled.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	34	100.0	153	4	US-09-228-986-83
2	33	97.1	457	4	US-09-252-991A-24232
3	32	94.1	196	4	US-09-252-991A-24657
4	32	94.1	308	4	US-09-252-991A-31991
5	32	94.1	367	3	US-08-605-150A-6
6	32	94.1	372	3	US-08-605-150A-6
7	31	91.2	227	3	US-08-244-646-17
8	31	91.2	342	1	US-08-244-646-15
9	31	91.2	342	1	US-08-592-936B-21
10	31	91.2	342	2	US-09-111-573-21
11	31	91.2	1196	3	US-08-881-706-2
12	30	88.2	354	2	US-08-948-569A-12
13	30	88.2	354	2	US-09-188-469-12
14	30	88.2	564	3	US-09-397-238A-12
15	30	88.2	582	4	US-09-252-991A-24081
16	29	85.3	90	4	US-09-328-986-92
17	29	85.3	116	4	US-09-352-991A-29189
18	29	85.3	139	4	US-09-352-991A-29189
19	29	85.3	164	4	US-09-585-173B-32
20	29	85.3	170	4	US-09-252-991A-24538
21	29	85.3	173	4	US-09-252-991A-21784
22	29	85.3	184	4	US-09-252-991A-18988
23	29	85.3	185	4	US-09-252-991A-18988
24	29	85.3	207	4	US-09-252-991A-23179
25	29	85.3	254	4	US-09-252-991A-23122
26	29	85.3	257	4	US-09-252-991A-25576
27	29	85.3	282	4	US-09-252-991A-18172
					Sequence 19978, A

28	29	85.3	287	4	US-09-252-991A-21133	Sequence 21133, A
29	29	85.3	390	4	US-09-252-991A-24517	Sequence 24517, A
30	29	85.3	396	4	US-09-252-991A-24697	Sequence 24697, A
31	29	85.3	397	4	US-09-252-991A-18709	Sequence 18709, A
32	29	85.3	410	1	US-08-123-343A-5	Sequence 5, App1
33	29	85.3	410	1	US-08-123-343A-5	Sequence 7, App1
34	29	85.3	410	1	US-09-431-573-5	Sequence 4, App1
35	29	85.3	410	3	US-09-431-573-5	Sequence 5, App1
36	29	85.3	421	4	US-09-252-991A-25017	Sequence 25017, A
37	29	85.3	421	4	US-09-252-991A-30742	Sequence 30742, A
38	29	85.3	481	4	US-09-252-991A-31018	Sequence 31018, A
39	29	85.3	523	2	US-08-473-553A-3	Sequence 3, App1
40	29	85.3	546	4	US-09-252-991A-25851	Sequence 25851, A
41	29	85.3	788	4	US-09-252-991A-28544	Sequence 28544, A
42	29	85.3	910	4	US-09-228-986-72	Sequence 72, App1
43	29	85.3	980	2	US-08-473-553A-6	Sequence 6, App1
44	29	85.3	985	2	US-08-473-553A-2	Sequence 2, App1
45	29	85.3	988	4	US-09-252-991A-29699	Sequence 29699, A

ALIGNMENTS

```

RESULT 1
US-09-228-986-83
; Sequence 83, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228, 986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-83

Query Match      100.0%  Score 34; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGRIP 6
Db      126 TGRIP 131

RESULT 2
US-09-252-991A-24232
; Sequence 24232, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24232
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24232

```

Query Match 97.1%; Score 33; DB 4; Length 457;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||:|
DB 32 TGRVPP 37

RESULT 3

US-09-252-991A-26457
; Sequence 26457, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26457
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26457

Query Match 94.1%; Score 32; DB 4; Length 196;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||:|
DB 34 TGRIPP 39

RESULT 4

US-09-252-991A-31991
; Sequence 31991, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31991
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31991

Query Match 94.1%; Score 32; DB 4; Length 308;
Best Local Similarity 83.3%; Pred. No. 81;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||:|
DB 28 TGRIPP 33

RESULT 5

US-08-605-150A-6
; Sequence 6, Application US/08605150A
; Patent No. 6103520

; GENERAL INFORMATION:
; APPLICANT: Topfer, Reinhard
; APPLICANT: Hausmann, Juedger
; APPLICANT: Schell, Jozef
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klein & Szekeres
; STREET: 4199 Campus Drive, Suite 700
; CITY: Irvine
; STATE: CA
; COUNTRY: USA
; ZIP: 92715

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/605,150A
APPLICATION NUMBER: US/08/605,150A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02936
FILING DATE: 02-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P4329827.3
FILING DATE: 03-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Szekeres, Gabor L.
REGISTRATION NUMBER: 28, 675
REFERENCE/DOCKET NUMBER: 542-04-PA

TELEPHONE: 714-854-4897
TELEFAX: 714-854-5502
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 367 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-605-150A-6

Query Match 94.1%; Score 32; DB 3; Length 367;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||:|
DB 347 TGRIPP 352

RESULT 6

US-08-605-150A-2
; Sequence 2, Application US/08605150A
; Patent No. 6103520
; GENERAL INFORMATION:
; APPLICANT: Topfer, Reinhard
; APPLICANT: Hausmann, Juedger
; APPLICANT: Schell, Jozef
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klein & Szekeres
; STREET: 4199 Campus Drive, Suite 700
; CITY: Irvine
; STATE: CA
; COUNTRY: USA

ZIP: 92715
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,150A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02936
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P4329827.3
FILING DATE: 03-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Szekeres, Gabor L.
REGISTRATION NUMBER: 28,675
REFERENCE/DOCKET NUMBER: 542-04-PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-854-5502
TELEFAX: 714-854-4897
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-150A-2

Query Match 94.1%; Score 32; DB 3; Length 372;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
|||:|
Db 352 TGRIP 357

RESULT 7
US-08-244-646-17
Sequence 17, Application US/08244646
Patent No. 5744692
GENERAL INFORMATION:
APPLICANT: Cervone, Felice
APPLICANT: De Lorenzo, Giulia
APPLICANT: Salvi, Giovanni
APPLICANT: Albersheim, Peter
APPLICANT: Darvill, Alan
APPLICANT: Bergmann, Carl
TITLE OF INVENTION: Nucleotide Sequences Coding An
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991

ADDRESS: Sally A. Sullivan
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/IT/00158
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 19-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-244-646-17

Query Match 91.2%; Score 31; DB 1; Length 227;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
|||:|
Db 101 TGRIP 106

RESULT 8
US-08-244-646-15
Sequence 15, Application US/08244646
Patent No. 5744692
GENERAL INFORMATION:
APPLICANT: Cervone, Felice
APPLICANT: De Lorenzo, Giulia
APPLICANT: Salvi, Giovanni
APPLICANT: Albersheim, Peter
APPLICANT: Darvill, Alan
APPLICANT: Bergmann, Carl
TITLE OF INVENTION: Nucleotide Sequences Coding An
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,646
FILING DATE: 06-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM 91A 000915
FILING DATE: 06-DEC-1991

ADDRESS: Sally A. Sullivan
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-244-646-15

Query Match
Best Local Similarity 91.2%; Score 31; DB 1; Length 342;
83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
11:1111
DB 216 TGRIPP 221

RESULT 9
US-08-592-936B-21
Sequence 21, Application US/08592936B
Patent No. 5783393
GENERAL INFORMATION:
APPLICANT: Kellogg, Jill A.
TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR
TITLE OF INVENTION: REGULATED EXPRESSION OF TRANSGENES IN PLANTS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,936B
FILING DATE: 29-JAN-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Susan T.
REGISTRATION NUMBER: 38,443
REFERENCE/DOCKET NUMBER: 4257-0012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Predicted amino acid coding sequence
INDIVIDUAL ISOLATE: of SEQ ID NO:20
US-08-592-936B-21

Query Match
Best Local Similarity 91.2%; Score 31; DB 1; Length 342;
83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
11:1111
DB 216 TGRIPP 221

RESULT 10
US-09-111-573-21
Sequence 21, Application US/09111573
Patent No. 5929302
GENERAL INFORMATION:

APPLICANT: Kellogg, Jill A.
APPLICANT: Bestwick, Richard K.
TITLE OF INVENTION: PLANT TISSUE / STAGE SPECIFIC PROMOTERS FOR
TITLE OF INVENTION: REGULATED EXPRESSION OF TRANSGENES IN PLANTS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,573
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,936
FILING DATE: 29-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Susan T.
REGISTRATION NUMBER: 38,443
REFERENCE/DOCKET NUMBER: 4257-0012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Predicted amino acid coding sequence
INDIVIDUAL ISOLATE: of SEQ ID NO:20
US-09-111-573-21

Query Match
Best Local Similarity 91.2%; Score 31; DB 2; Length 342;
83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
11:1111
DB 216 TGRIPP 221

RESULT 11
US-08-881-706-2
Sequence 2, Application US/08881706
Patent No. 6245969
GENERAL INFORMATION:
APPLICANT: Chong, Joane
APPLICANT: Li, Jianming
TITLE OF INVENTION: Receptor Kinase BIN1
FILE REFERENCE: 07251/022001
CURRENT APPLICATION NUMBER: US/08/881,706
CURRENT FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1196
TYPE: PRT
ORGANISM: Arabidopsis sp.
US-08-881-706-2

Query Match
Best Local Similarity 91.2%; Score 31; DB 3; Length 1196;
83.3%; Pred. No. 4.7e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGRIPP 6
11:111
DB 429 TGRIPP 434

RESULT 12

US-08-948-569A-12
; Sequence 12, Application US/08948569A
; Patent No. 5882926
; GENERAL INFORMATION:
; APPLICANT: Amara, Susan G
; APPLICANT: Arriza, Jeffrey L
; APPLICANT: Eliasof, Scott
; APPLICANT: Kavanaugh, Michael P
; TITLE OF INVENTION: Excitatory Amino Acid Transporter Genes
; TITLE OF INVENTION: and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,569A
; FILING DATE: 10-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5882926nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,509-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-948-569A-12

Query Match 88.2%; Score 30; DB 2; Length 564;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGRIPP 6
11:111
DB 171 TGRIPP 176

RESULT 13

US-09-188-469-12
; Sequence 12, Application US/09188469
; Patent No. 5989825
; GENERAL INFORMATION:
; APPLICANT: Amara, Susan G
; APPLICANT: Arriza, Jeffrey L
; APPLICANT: Eliasof, Scott
; APPLICANT: Kavanaugh, Michael P
; TITLE OF INVENTION: Excitatory Amino Acid Transporter Genes
; TITLE OF INVENTION: and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,469
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,569
; FILING DATE: 10-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5989825nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,509-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-188-469-12

Query Match 88.2%; Score 30; DB 2; Length 564;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGRIPP 6
11:111
DB 171 TGRIPP 176

RESULT 14

US-09-397-238A-12
; Sequence 12, Application US/09397238A
; Patent No. 6284505
; GENERAL INFORMATION:
; APPLICANT: Amara, Susan G
; APPLICANT: Arriza, Jeffrey L
; APPLICANT: Eliasof, Scott
; APPLICANT: Kavanaugh, Michael P
; TITLE OF INVENTION: Excitatory Amino Acid Transporter Genes
; TITLE OF INVENTION: and Uses
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/397,238A
; FILING DATE: 16-Sep-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6284505nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,509-F

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 564 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-09-397-238A-12

Query Match 88.2%; Score 30; DB 3; Length 564;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
 ||:|
 Db 171 TGRVPP 176

RESULT 15
 US-09-252-991A-20481
 ; Sequence 20481, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20481
 ; LENGTH: 582
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (388)
 ; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
 ; US-09-252-991A-20481

Query Match 88.2%; Score 30; DB 4; Length 582;
 Best Local Similarity 83.3%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGRIPP 6
 |||
 Db 94 TGRFPP 99

Search completed: August 20, 2003, 12:44:23
 Job time: 4.15663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 6.50602 Seconds

(without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIP 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCMS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	6	11	US-09-300-425B-34
2	34	100.0	108	11	US-09-300-425B-21
3	34	100.0	109	15	US-10-101-464A-685
4	34	100.0	153	15	US-10-101-464A-83
5	34	100.0	155	15	US-10-101-464A-645
6	34	100.0	163	15	US-10-101-464A-766
7	34	100.0	224	15	US-10-101-464A-751
8	34	100.0	705	15	US-10-101-464A-894
9	32	94.1	598	15	US-10-156-761-8911
10	31	91.2	147	15	US-10-101-464A-701
11	31	91.2	383	15	US-10-101-464A-898
12	31	91.2	1196	11	US-09-823-394-2
13	30	88.2	63	9	US-09-864-761-33825
14	30	88.2	370	15	US-10-101-464A-944
15	30	88.2	604	10	US-09-738-269-14

16	30	88.2	864	15	US-10-101-464A-896	Sequence 896, App
17	29	85.3	90	15	US-10-101-464A-92	Sequence 92, App
18	29	85.3	132	15	US-10-101-464A-572	Sequence 572, App
19	29	85.3	161	15	US-10-101-464A-570	Sequence 570, App
20	29	85.3	188	16	US-10-080-170-290	Sequence 290, App
21	29	85.3	215	15	US-10-156-761-14238	Sequence 14238, A
22	29	85.3	226	15	US-10-156-761-12101	Sequence 12101, A
23	29	85.3	277	15	US-10-101-464A-631	Sequence 631, App
24	29	85.3	281	15	US-10-101-464A-619	Sequence 619, App
25	29	85.3	304	11	US-09-764-891-4037	Sequence 4037, App
26	29	85.3	410	14	US-10-108-605-89	Sequence 89, App
27	29	85.3	418	9	US-09-816-664-2	Sequence 2, App
28	29	85.3	478	10	US-09-738-626-3883	Sequence 3883, App
29	29	85.3	514	12	US-10-216-163-6	Sequence 6, App
30	29	85.3	514	15	US-10-227-884-6	Sequence 6, App
31	29	85.3	514	15	US-10-230-163-6	Sequence 6, App
32	29	85.3	514	15	US-10-230-338-6	Sequence 6, App
33	29	85.3	514	15	US-10-218-631-6	Sequence 6, App
34	29	85.3	514	15	US-10-218-414-6	Sequence 6, App
35	29	85.3	514	15	US-10-216-159A-6	Sequence 6, App
36	29	85.3	514	15	US-10-218-849-6	Sequence 6, App
37	29	85.3	514	15	US-10-227-883-6	Sequence 6, App
38	29	85.3	514	15	US-10-227-883-6	Sequence 6, App
39	29	85.3	514	15	US-10-219-076-6	Sequence 6, App
40	29	85.3	514	15	US-10-230-434-6	Sequence 6, App
41	29	85.3	514	15	US-10-219-003-6	Sequence 6, App
42	29	85.3	514	15	US-10-219-075-6	Sequence 6, App
43	29	85.3	514	15	US-10-219-464-6	Sequence 6, App
44	29	85.3	514	15	US-10-219-466-6	Sequence 6, App
45	29	85.3	514	15	US-10-219-479-6	Sequence 6, App

ALIGNMENTS

RESULT 1
US-09-300-425B-34
; Sequence 34, Application US/09300425B
; Publication No. US20030045681A1
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
; OTHER INFORMATION: antibody clone
US-09-300-425B-34

Query Match 100.0%; Score 34; DB 11; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
Db 1 TGRIP 6

RESULT 2

US-09-300-425B-21
; Sequence 21, Application US/09300425B
; Publication No. US20030045681A1
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VL antibody
US-09-300-425B-21

Query Match 100.0%; Score 34; DB 11; Length 108;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||||
DB 92 TGRIPP 97

RESULT 3
US-10-101-464A-685
; Sequence 685, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-685

Query Match 100.0%; Score 34; DB 15; Length 109;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||||
DB 21 TGRIPP 26

RESULT 4
US-10-101-464A-83
; Sequence 83, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-83

Query Match 100.0%; Score 34; DB 15; Length 153;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||||
DB 126 TGRIPP 131

RESULT 5
US-10-101-464A-645
; Sequence 645, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 645
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-645

Query Match 100.0%; Score 34; DB 15; Length 155;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6

Db 116 TGRIP 121

RESULT 6
US-10-101-464A-766

; Sequence 766, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 766
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-766

Query Match 100.0%; Score 34; DB 15; Length 163;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
Db 126 TGRIP 131

RESULT 7

; Sequence 751, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 751
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-751

Query Match 100.0%; Score 34; DB 15; Length 224;
Best Local Similarity 100.0%; Pred. No. 58;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
Db 188 TGRIP 193

RESULT 8

; Sequence 894, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-894

Query Match 100.0%; Score 34; DB 15; Length 705;
Best Local Similarity 100.0%; Pred. No. 17e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
Db 126 TGRIP 131

RESULT 9

; Sequence 891, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8911
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8911

Query Match 94.1%; Score 32; DB 15; Length 598;

Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
11:111
Db 356 TGRIPP 361

RESULT 10

US-10-101-464A-701
; Sequence 701, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 701
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-701

Query Match 91.2%; Score 31; DB 15; Length 147;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
11:111
Db 27 TGRIPP 32

RESULT 11

US-10-101-464A-898
; Sequence 898, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 898
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Pinus radiata

US-10-101-464A-898

Query Match 91.2%; Score 31; DB 15; Length 383;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
11:111
Db 208 TGRIPP 213

RESULT 12

US-09-823-394-2
; Sequence 2, Application US/09823394
; Publication No. US20030041344A1
; GENERAL INFORMATION:
; APPLICANT: Chory, Joanne
; APPLICANT: Jianming, Li
; APPLICANT: Salk Institute for Biological Studies
; TITLE OF INVENTION: RECEPTOR KINASE, BIN 1
; FILE REFERENCE: SALKINS.012CP1
; CURRENT APPLICATION NUMBER: US/09/823,394
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 08/881,706
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1196
; TYPE: PRT
; ORGANISM: Arabidopsis
US-09-823-394-2

Query Match 91.2%; Score 31; DB 11; Length 1196;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
11:111
Db 429 TGRIPP 434

RESULT 13

US-09-864-761-33825
; Sequence 33825, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 33825
;; LENGTH: 63
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007900.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8.4
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;; OTHER INFORMATION: EST_HUMAN HIT: A1243801.1, EVALUE 8.60e+00
;; OTHER INFORMATION: SWISSPROT HIT: Q13085, EVALUE 1.00e-30
US-09-864-761-33825

Query Match      88.2%; Score 30; DB 9; Length 63;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGRIP 6
        :|||||
Db      30 SGRIP 35

RESULT 14
US-10-101-464A-944
;; Sequence 944, Application US/10101464A
;; Publication No. US20030046728A1
;; GENERAL INFORMATION:
;; APPLICANT: Strabala, Timothy
;; APPLICANT: Mieuwenhuizen, Nicolaas
;; APPLICANT: Higgins, Colleen M.
;; TITLE OF INVENTION: Compositions isolated from Plant Cells
;; TITLE OF INVENTION: and their use in the modification of Plant Cell Signaling
;; FILE REFERENCE: 11000.1020c2
;; CURRENT APPLICATION NUMBER: US/10/101,464A
;; CURRENT FILING DATE: 2002-03-18
;; PRIOR APPLICATION NUMBER: 09/704,302
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 09/228,986
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/162,866
;; PRIOR FILING DATE: 1999-11-01
;; PRIOR APPLICATION NUMBER: PCT/US00/00724
;; PRIOR FILING DATE: 2000-01-11
;; NUMBER OF SEQ ID NOS: 989
;; SOFTWARE: FastSeq for Windows Version 4.0
```

```
;; SEQ ID NO 944
;; LENGTH: 370
;; TYPE: PRT
;; ORGANISM: Pinus radiata
US-10-101-464A-944

Query Match      88.2%; Score 30; DB 15; Length 370;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGRIP 6
        :|||||
Db      354 TGRIP 359

RESULT 15
US-09-758-269-14
;; Sequence 14, Application US/09758269
;; Patent No. US20020104120A1
;; GENERAL INFORMATION:
;; APPLICANT: IUCHI, SATOSHI
;; APPLICANT: KOBAYASHI, MASATOMO
;; APPLICANT: SHINOZAKI, KAZUO
;; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
;; FILE REFERENCE: 3914-3
;; CURRENT APPLICATION NUMBER: US/09/758,269
;; CURRENT FILING DATE: 2001-01-12
;; PRIOR APPLICATION NUMBER: JP 2001-003476
;; PRIOR FILING DATE: 2001-01-11
;; PRIOR APPLICATION NUMBER: JP 2000-010056
;; PRIOR FILING DATE: 2000-01-13
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn ver. 2.1
;; SEQ ID NO 14
;; LENGTH: 604
;; TYPE: PRT
;; ORGANISM: Zea mays
US-09-758-269-14

Query Match      88.2%; Score 30; DB 10; Length 604;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGRIP 6
        :|||||
Db      142 SGRIP 147
```

Search completed: August 20, 2003, 13:16:48
Job time : 7.50602 secs

THIS PAGE BLANK (USPTO)

	Matches	6; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1	TGRIP 6			
Db	1	TGRIP 6			

```

RESULT 2
US-09-300-425B-34
Sequence 34, Application US/09300425B
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARI, Lorenzo
APPLICANT: VITI, Francesca
APPLICANT: BIRCHLER, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
TITLE OF INVENTION: ANGIOGENESIS
FILE REFERENCE: SCH-1733P1
CURRENT APPLICATION NUMBER: US/09/300,425B
CURRENT FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 09/075,338
PRIOR FILING DATE: 1998-05-11
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
US-09-300-425B-34

```

	Query Match	100.0%;	Score 34;	DB 17;	Length 6;
	Best Local Similarity	100.0%;	Pred. No.	5.2e+05;	
Matches	6;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1 TGRIP 6				
DB	1 TGRIP 6				

```

      RESULT 3
US-09-512-082-34
; Sequence 34, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITTI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTILLARPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-173P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anti-ED-B
; OTHER INFORMATION: antibody clone
US-09-512-082-34

```

query Match	100.0%; Score 34; DB 19; Length 6;
-------------	------------------------------------

Best Local Similarity 100.0%; Pred. NO. 5.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TGRIPP	6
Db	1	TGRIPP	6

```

RESULT 4
US-09-075-338C-21
: Sequence 21, Application US/09075338C
:
: GENERAL INFORMATION:
: APPLICANT: NERI, Dario
: APPLICANT: TARI, Lorenzo
: APPLICANT: VITI, Francesca
: APPLICANT: BIRCHLER, Manfred
: TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
: FILE REFERENCE: SCH-1733
: CURRENT APPLICATION NUMBER: US/09/075.338C
: CURRENT FILING DATE: 1998-05-11
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
:
: LENGTH: 108
:
: TYPE: PRT
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: VL antibody
: OTHER INFORMATION: Specific for Ed-B domain of fibronectin
US-09-075-338C-21

```

Query Match	100.0%;	Score 34;	DB 14;	Length 108;
Best Local Similarity	100.0%;	Pred. No. 4.8e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
				Indels 0;
QY	1	TGRIP 6		
	111111			
DB	92	TGRIP 97		

```

1      RESULT 5
2      US-09-300-425B-21
3      ; Sequence 21, Application US/09300425B
4      ; GENERAL INFORMATION:
5      ; APPLICANT: NERI, Dario
6      ; APPLICANT: TARLI, Lorenzo
7      ; APPLICANT: VITTI, Francesca
8      ; APPLICANT: BIRCHLER, Manfred
9      ; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONJUGATES
10     ; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
11     ; TITLE OF INVENTION: ANGIOGENESIS
12     ; FILE REFERENCE: SCH-1733P1
13     ; CURRENT APPLICATION NUMBER: US/09/300,425B
14     ; CURRENT FILING DATE: 1999-04-28
15     ; PRIOR APPLICATION NUMBER: 09/075,338
16     ; PRIOR FILING DATE: 1998-05-11
17     ; NUMBER OF SEQ ID NOS: 34
18     ; SOFTWARE: PatentIn Ver. 2.1
19     ; SEQ ID NO 21
20     ;     ;
21     ;     ;
22     ;     ;
23     ;     ;
24     ;     ;
25     ;     ;
26     ;     ;
27     ;     ;
28     ;     ;
29     ;     ;
30     ;     ;
31     ;     ;
32     ;     ;
33     ;     ;
34     ;     ;
35     ;     ;
36     ;     ;
37     ;     ;
38     ;     ;
39     ;     ;
40     ;     ;
41     ;     ;
42     ;     ;
43     ;     ;
44     ;     ;
45     ;     ;
46     ;     ;
47     ;     ;
48     ;     ;
49     ;     ;
50     ;     ;
51     ;     ;
52     ;     ;
53     ;     ;
54     ;     ;
55     ;     ;
56     ;     ;
57     ;     ;
58     ;     ;
59     ;     ;
60     ;     ;
61     ;     ;
62     ;     ;
63     ;     ;
64     ;     ;
65     ;     ;
66     ;     ;
67     ;     ;
68     ;     ;
69     ;     ;
70     ;     ;
71     ;     ;
72     ;     ;
73     ;     ;
74     ;     ;
75     ;     ;
76     ;     ;
77     ;     ;
78     ;     ;
79     ;     ;
80     ;     ;
81     ;     ;
82     ;     ;
83     ;     ;
84     ;     ;
85     ;     ;
86     ;     ;
87     ;     ;
88     ;     ;
89     ;     ;
90     ;     ;
91     ;     ;
92     ;     ;
93     ;     ;
94     ;     ;
95     ;     ;
96     ;     ;
97     ;     ;
98     ;     ;
99     ;     ;
100    ;     ;
101    ;     ;
102    ;     ;
103    ;     ;
104    ;     ;
105    ;     ;
106    ;     ;
107    ;     ;
108    ;     ;
109    ;     ;
110    ;     ;
111    ;     ;
112    ;     ;
113    ;     ;
114    ;     ;
115    ;     ;
116    ;     ;
117    ;     ;
118    ;     ;
119    ;     ;
120    ;     ;
121    ;     ;
122    ;     ;
123    ;     ;
124    ;     ;
125    ;     ;
126    ;     ;
127    ;     ;
128    ;     ;
129    ;     ;
130    ;     ;
131    ;     ;
132    ;     ;
133    ;     ;
134    ;     ;
135    ;     ;
136    ;     ;
137    ;     ;
138    ;     ;
139    ;     ;
140    ;     ;
141    ;     ;
142    ;     ;
143    ;     ;
144    ;     ;
145    ;     ;
146    ;     ;
147    ;     ;
148    ;     ;
149    ;     ;
150    ;     ;
151    ;     ;
152    ;     ;
153    ;     ;
154    ;     ;
155    ;     ;
156    ;     ;
157    ;     ;
158    ;     ;
159    ;     ;
160    ;     ;
161    ;     ;
162    ;     ;
163    ;     ;
164    ;     ;
165    ;     ;
166    ;     ;
167    ;     ;
168    ;     ;
169    ;     ;
170    ;     ;
171    ;     ;
172    ;     ;
173    ;     ;
174    ;     ;
175    ;     ;
176    ;     ;
177    ;     ;
178    ;     ;
179    ;     ;
180    ;     ;
181    ;     ;
182    ;     ;
183    ;     ;
184    ;     ;
185    ;     ;
186    ;     ;
187    ;     ;
188    ;     ;
189    ;     ;
190    ;     ;
191    ;     ;
192    ;     ;
193    ;     ;
194    ;     ;
195    ;     ;
196    ;     ;
197    ;     ;
198    ;     ;
199    ;     ;
200    ;     ;
201    ;     ;
202    ;     ;
203    ;     ;
204    ;     ;
205    ;     ;
206    ;     ;
207    ;     ;
208    ;     ;
209    ;     ;
210    ;     ;
211    ;     ;
212    ;     ;
213    ;     ;
214    ;     ;
215    ;     ;
216    ;     ;
217    ;     ;
218    ;     ;
219    ;     ;
220    ;     ;
221    ;     ;
222    ;     ;
223    ;     ;
224    ;     ;
225    ;     ;
226    ;     ;
227    ;     ;
228    ;     ;
229    ;     ;
230    ;     ;
231    ;     ;
232    ;     ;
233    ;     ;
234    ;     ;
235    ;     ;
236    ;     ;
237    ;     ;
238    ;     ;
239    ;     ;
240    ;     ;
241    ;     ;
242    ;     ;
243    ;     ;
244    ;     ;
245    ;     ;
246    ;     ;
247    ;     ;
248    ;     ;
249    ;     ;
250    ;     ;
251    ;     ;
252    ;     ;
253    ;     ;
254    ;     ;
255    ;     ;
256    ;     ;
257    ;     ;
258    ;     ;
259    ;     ;
260    ;     ;
261    ;     ;
262    ;     ;
263    ;     ;
264    ;     ;
265    ;     ;
266    ;     ;
267    ;     ;
268    ;     ;
269    ;     ;
270    ;     ;
271    ;     ;
272    ;     ;
273    ;     ;
274    ;     ;
275    ;     ;
276    ;     ;
277    ;     ;
278    ;     ;
279    ;     ;
280    ;     ;
281    ;     ;
282    ;     ;
283    ;     ;
284    ;     ;
285    ;     ;
286    ;     ;
287    ;     ;
288    ;     ;
289    ;     ;
290    ;     ;
291    ;     ;
292    ;     ;
293    ;     ;
294    ;     ;
295    ;     ;
296    ;     ;
297    ;     ;
298    ;     ;
299    ;     ;
300    ;     ;
301    ;     ;
302    ;     ;
303    ;     ;
304    ;     ;
305    ;     ;
306    ;     ;
307    ;     ;
308    ;     ;
309    ;     ;
310    ;     ;
311    ;     ;
312    ;     ;
313    ;     ;
314    ;     ;
315    ;     ;
316    ;     ;
317    ;     ;
318    ;     ;
319    ;     ;
320    ;     ;
321    ;     ;
322    ;     ;
323    ;     ;
324    ;     ;
325    ;     ;
326    ;     ;
327    ;     ;
328    ;     ;
329    ;     ;
330    ;     ;
331    ;     ;
332    ;     ;
333    ;     ;
334    ;     ;
335    ;     ;
336    ;     ;
337    ;     ;
338    ;     ;
339    ;     ;
340    ;     ;
341    ;     ;
342    ;     ;
343    ;     ;
344    ;     ;
345    ;     ;
346    ;     ;
347    ;     ;
348    ;     ;
349    ;     ;
350    ;     ;
351    ;     ;
352    ;     ;
353    ;     ;
354    ;     ;
355    ;     ;
356    ;     ;
357    ;     ;
358    ;     ;
359    ;     ;
360    ;     ;
361    ;     ;
362    ;     ;
363    ;     ;
364    ;     ;
365    ;     ;
366    ;     ;
367    ;     ;
368    ;     ;
369    ;     ;
370    ;     ;
371    ;     ;
372    ;     ;
373    ;     ;
374    ;     ;
375    ;     ;
376    ;     ;
377    ;     ;
378    ;     ;
379    ;     ;
380    ;     ;
381    ;     ;
382    ;     ;
383    ;     ;
384    ;     ;
385    ;     ;
386    ;     ;
387    ;     ;
388    ;     ;
389    ;     ;
390    ;     ;
391    ;     ;
392    ;     ;
393    ;     ;
394    ;     ;
395    ;     ;
396    ;     ;
397    ;     ;
398    ;     ;
399    ;     ;
400    ;     ;
401    ;     ;
402    ;     ;
403    ;     ;
404    ;     ;
405    ;     ;
406    ;     ;
407    ;     ;
408    ;     ;
409    ;     ;
410    ;     ;
411    ;     ;
412    ;     ;
413    ;     ;
414    ;     ;
415    ;     ;
416    ;     ;
417    ;     ;
418    ;     ;
419    ;     ;
420    ;     ;
421    ;     ;
422    ;     ;
423    ;     ;
424    ;     ;
425    ;     ;
426    ;     ;
427    ;     ;
428    ;     ;
429    ;     ;
430    ;     ;
431    ;     ;
432    ;     ;
433    ;     ;
434    ;     ;
435    ;     ;
436    ;     ;
437    ;     ;
438    ;     ;
439    ;     ;
440    ;     ;
441    ;     ;
442    ;     ;
443    ;     ;
444    ;     ;
445    ;     ;
446    ;     ;
447    ;     ;
448    ;     ;
449    ;     ;
450    ;     ;
451    ;     ;
452    ;     ;
453    ;     ;
454    ;     ;
455    ;     ;
456    ;     ;
457    ;     ;
458    ;     ;
459    ;     ;
460    ;     ;
461    ;     ;
462    ;     ;
463    ;     ;
464    ;     ;
465    ;     ;
466    ;     ;
467    ;     ;
468    ;     ;
469    ;     ;
470    ;     ;
471    ;     ;
472    ;     ;
473    ;     ;
474    ;     ;
475    ;     ;
476    ;     ;
477    ;     ;
478    ;     ;
479    ;     ;
480    ;     ;
481    ;     ;
482    ;     ;
483    ;     ;
484    ;     ;
485    ;     ;
486    ;     ;
487    ;     ;
488    ;     ;
489    ;     ;
490    ;     ;
491    ;     ;
492    ;     ;
493    ;     ;
494    ;     ;
495    ;     ;
496    ;     ;
497    ;     ;
498    ;     ;
499    ;     ;
500    ;     ;
501    ;
```

Query Match	100.0%;	Score 34;	DB 17;	length 108;
Best Local	Similarity	100.0%;	Pred. No. 4.Be+02;	
Matches	6;	Conservative	0;	Indels 0; Gaps 0;
		Mismatches	0;	

QY 1 TGRIP 6
|||||

Db 92 TGRIPP 97

RESULT 6
US-09-512-082-21

; Sequence 21, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARTI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHLER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VL antibody
; US-09-512-082-21

Query Match 100.0%; Score 34; DB 19; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 92 TGRIPP 97

RESULT 7
PCT-US00-00724-685
; Sequence 685, Application PC/TUS0000724
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C1PCT
; CURRENT APPLICATION NUMBER: PCT/US00/00724
; CURRENT FILING DATE: 2000-01-11
; EARLIER APPLICATION NUMBER: US 09/228,986
; EARLIER FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: US 60/162,866
; EARLIER FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 1322
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 685
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pinus radiata
; PCT-US00-00724-685

Query Match 100.0%; Score 34; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 21 TGRIPP 26RESULT 8
US-09-704-302A-685

; Sequence 685, Application US/09704302A
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C1U
; CURRENT APPLICATION NUMBER: US/09/704,302A
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 1402
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pinus radiata
; US-09-704-302A-685

Query Match 100.0%; Score 34; DB 21; Length 109;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 21 TGRIPP 26

RESULT 9
US-10-101-464A-685
; Sequence 685, Application US/10101464A
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 685
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pinus radiata
; US-10-101-464A-685

Query Match 100.0%; Score 34; DB 27; Length 109;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 21 TGRIPP 26

RESULT 10
US-60-162-866-685
; Sequence 685, Application US/60162866
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling

```

; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/60/162,866
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 1275
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 685
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pinus radiata
US-60-162-866-685

```

```

Query Match          100.0%; Score 34; DB 31; Length 109;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGRIPP 6
        |||||
Db      21 TGRIPP 26

```

```

RESULT 11
US-10-424-599-167644
; Sequence 167644, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167644
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(114)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12239C.1.pep
US-10-424-599-167644

```

```

Query Match          100.0%; Score 34; DB 30; Length 114;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGRIPP 6
        |||||
Db      51 TGRIPP 56

```

```

RESULT 12
US-09-513-999C-4223
; Sequence 4223, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4223
; LENGTH: 144
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -27...-1
; OTHER INFORMATION: score 8.6
; OTHER INFORMATION: seq ATVLLSPGSVAA/SH
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 108
; OTHER INFORMATION: Xaa-Ile or Leu or Val
US-09-513-999C-4223

```

```

Query Match          100.0%; Score 34; DB 19; Length 144;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGRIPP 6
        |||||
Db      136 TGRIPP 141

```

```

RESULT 13
PCT-US00-00724-83
; Sequence 83, Application PC/TUS0000724
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020C1PCT
; CURRENT APPLICATION NUMBER: PCT/US00/00724
; CURRENT FILING DATE: 2000-01-11
; EARLIER APPLICATION NUMBER: US 09/228,986
; EARLIER FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: US 60/162,866
; EARLIER FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 1322
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
PCT-US00-00724-83

```

```

Query Match          100.0%; Score 34; DB 1; Length 153;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGRIPP 6
        |||||
Db      126 TGRIPP 131

```

```

RESULT 14
US-09-704-302A-83
; Sequence 83, Application US/09704302A
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells.
; FILE REFERENCE: 11000.1020C1U
; CURRENT APPLICATION NUMBER: US/09/704,302A
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 1402
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-704-302A-83

```

```

Query Match          100.0%; Score 34; DB 21; Length 153;

```

Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||||

DB 126 TGRIPP 131

RESULT 15

US-10-101-464A-83

; Sequence 83, Application US/10101464A

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions isolated from plant cells

; TITLE OF INVENTION: and their use in the modification of plant cell signaling

; FILE REFERENCE: 11000.10202

; CURRENT APPLICATION NUMBER: US/10/101,464A

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 83

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Pinus radiata

US-10-101-464A-83

Query Match 100.0%; Score 34; DB 27; Length 153;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
|||||

DB 126 TGRIPP 131

Search completed: August 20, 2003, 13:13:46
Job time: 62.4096 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 12:34:01 ; Search time 1.73494 Seconds
(without alignments)
91.710 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIPP 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 126775 seqs, 26518662 residues

Total number of hits satisfying chosen parameters: 126775

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/pae/PCT_NEW.COMB.pep.*
2: /cgn2_6/ptodata/1/pae/US06_NEW.COMB.pep.*
3: /cgn2_6/ptodata/1/pae/US07_NEW.COMB.pep.*
4: /cgn2_6/ptodata/1/pae/US08_NEW.COMB.pep.*
5: /cgn2_6/ptodata/1/pae/US09_NEW.COMB.pep.*
6: /cgn2_6/ptodata/1/pae/US10_NEW.COMB.pep.*
7: /cgn2_6/ptodata/1/pae/US60_NEW.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	238	6	US-10-336-041A-1
2	34	100.0	240	6	US-10-336-041A-10
3	34	100.0	240	6	US-10-336-041A-12
4	34	100.0	241	6	US-10-336-041A-11
5	34	100.0	241	6	US-10-336-041A-13
6	34	100.0	247	6	US-10-336-041A-9
7	34	100.0	307	6	US-10-291-265-332
8	34	100.0	311	6	US-10-408-765A-1931
9	34	100.0	312	6	US-10-291-265-334
10	33	97.1	518	6	US-10-612-783-4450
11	32	94.1	639	6	US-10-408-765A-2802
12	30	88.2	334	6	US-10-286-897-3022
13	30	88.2	334	6	US-10-258-898A-3022
14	30	88.2	365	6	US-10-258-897-6594
15	30	88.2	365	6	US-10-258-898A-6594
16	30	88.2	1109	6	US-10-613-520-1196
17	30	88.2	1109	6	US-10-613-520-2378
18	29	85.3	99	6	US-09-674-546A-1527
19	29	85.3	99	6	US-09-674-546A-1529
20	29	85.3	175	6	US-10-603-108-3445
21	29	85.3	191	1	PCT-US03-20480-18
22	29	85.3	216	1	PCT-US03-20480-20
23	29	85.3	224	1	PCT-US03-20480-15
24	29	85.3	267	6	US-10-293-244-3884
25	29	85.3	272	6	US-10-286-897-6954
26	29	85.3	272	6	US-10-258-898A-6554

27	29	85.3	338	1	PCT-US03-20480-16	Sequence 16, Appl
28	29	85.3	352	6	US-10-603-108-3272	Sequence 3272, Ap
29	29	85.3	358	6	US-10-286-897-3366	Sequence 3366, Ap
30	29	85.3	358	6	US-10-258-898A-3366	Sequence 3366, Ap
31	29	85.3	473	6	US-10-293-244-1916	Sequence 1916, Ap
32	29	85.3	484	6	US-10-293-244-3743	Sequence 3743, Ap
33	29	85.3	709	6	US-10-603-114-6114	Sequence 6114, Ap
34	29	85.3	746	6	US-10-293-244-1775	Sequence 1775, Ap
35	29	85.3	1336	6	US-10-408-765A-1449	Sequence 1449, Ap
36	29	85.3	2483	6	US-10-273-573-10255	Sequence 10255, A
37	28	82.4	11	6	US-10-601-837-73	Sequence 73, Appl
38	28	82.4	129	6	US-10-273-573-5774	Sequence 5774, Ap
39	28	82.4	131	6	US-10-293-244-3552	Sequence 3552, Ap
40	28	82.4	136	1	PCT-US03-23932-1	Sequence 1, Appl1
41	28	82.4	136	1	PCT-US03-23932-2	Sequence 2, Appl1
42	28	82.4	136	1	PCT-US03-23932-3	Sequence 3, Appl1
43	28	82.4	136	1	PCT-US03-23932-4	Sequence 4, Appl1
44	28	82.4	136	1	PCT-US03-23932-5	Sequence 5, Appl1
45	28	82.4	189	6	US-10-273-573-6504	Sequence 6504, Ap

ALIGNMENTS

RESULT 1
US-10-336-041A-1
Sequence 1, Application US/10336041A
GENERAL INFORMATION:
APPLICANT: Schering AG
TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
FILE REFERENCE: 27041P.WOAS
CURRENT APPLICATION NUMBER: US/10/336, 041A
CURRENT FILING DATE: 2003-01-03
PRIOR APPLICATION NUMBER: EP02 000 315.8
PRIOR FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: US60/358702
PRIOR FILING DATE: 2002-02-25
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 238
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(116)
OTHER INFORMATION: VH
FEATURE:
NAME/KEY: SITE
LOCATION: (117)..(130)
OTHER INFORMATION: Linker
FEATURE:
NAME/KEY: SITE
LOCATION: (131)..(238)
OTHER INFORMATION: VL
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant
OTHER INFORMATION: antibody fragment
US-10-336-041A-1

Query Match 100.0%; Score 34; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 222 TGRIPP 227

RESULT 2
US-10-336-041A-10
Sequence 10, Application US/10336041A
GENERAL INFORMATION:

APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P.WOAS
; CURRENT APPLICATION NUMBER: US/10/336,041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
; OTHER INFORMATION: antibody fragment
US-10-336-041A-10

Query Match
Best Local Similarity 100.0%; Score 34; DB 6; Length 240;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
|||||
Db 220 TGRIPP 225

RESULT 3
US-10-336-041A-12
; Sequence 12, Application US/10336041A
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P.WOAS
; CURRENT APPLICATION NUMBER: US/10/336,041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
; OTHER INFORMATION: antibody fragment
US-10-336-041A-12

Query Match
Best Local Similarity 100.0%; Score 34; DB 6; Length 240;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
|||||
Db 220 TGRIPP 225

RESULT 4
US-10-336-041A-11
; Sequence 11, Application US/10336041A
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P.WOAS
; CURRENT APPLICATION NUMBER: US/10/336,041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 247
; TYPE: PRT

PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
; OTHER INFORMATION: antibody fragment
US-10-336-041A-11

Query Match
Best Local Similarity 100.0%; Score 34; DB 6; Length 241;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
|||||
Db 220 TGRIPP 225

RESULT 5
US-10-336-041A-13
; Sequence 13, Application US/10336041A
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P.WOAS
; CURRENT APPLICATION NUMBER: US/10/336,041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
; OTHER INFORMATION: antibody fragment
US-10-336-041A-13

Query Match
Best Local Similarity 100.0%; Score 34; DB 6; Length 241;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
|||||
Db 220 TGRIPP 225

RESULT 6
US-10-336-041A-9
; Sequence 9, Application US/10336041A
; GENERAL INFORMATION:
; APPLICANT: Schering AG
; TITLE OF INVENTION: New methods for diagnosis and treatment of tumours
; FILE REFERENCE: 27041P.WOAS
; CURRENT APPLICATION NUMBER: US/10/336,041A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: EP02 000 315.8
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: US60/358702
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 247
; TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant
OTHER INFORMATION: antibody fragment
US-10-336-041A-9

Query Match 100.0%; Score 34; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 220 TGRIPP 225

RESULT 7
US-10-291-265-332
; Sequence 332, Application US/10291265
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 332
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-332

Query Match 100.0%; Score 34; DB 6; Length 307;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 241 TGRIPP 246

RESULT 8
US-10-408-765A-1931
; Sequence 1931, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wernock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1931
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1931

Query Match 100.0%; Score 34; DB 6; Length 311;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 52 TGRIPP 57

RESULT 9
US-10-291-265-334
; Sequence 334, Application US/10291265
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 334
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-334

Query Match 100.0%; Score 34; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
Db 246 TGRIPP 251

RESULT 10
US-10-612-783-4490
; Sequence 4490, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalig, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 4490
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRI4577_129850C.1.dep
US-10-612-783-4490

Query Match 97.1%; Score 33; DB 6; Length 518;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
|||||

Db 504 TGRVP 509

```
RESULT 11
US-10-408-765A-2802
; Sequence 2802, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Chosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2802
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2802
```

```
Query Match
Best Local Similarity 94.1%; Score 32; DB 6; Length 639;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGRIP 6
Db 618 TGRLP 623
```

```
RESULT 12
US-10-286-897-3022
; Sequence 3022, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 3022
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-3022
```

```
Query Match
Best Local Similarity 88.2%; Score 30; DB 6; Length 334;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 TGRIP 6
Db 274 TGRIP 279

```
RESULT 13
US-10-258-898A-3022
; Sequence 3022, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 3022
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-3022
```

```
Query Match
Best Local Similarity 88.2%; Score 30; DB 6; Length 334;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGRIP 6
Db 274 TGRIP 279
```

```
RESULT 14
US-10-286-897-6594
; Sequence 6594, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
```

; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 6594
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-6594

Query Match 88.2%; Score 30; DB 6; Length 365;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
||:||||
Db 305 TGRIPP 310

RESULT 15
US-10-258-898A-6594
; Sequence 6594, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 6594
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-6594

Query Match 88.2%; Score 30; DB 6; Length 365;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
||:||||
Db 305 TGRIPP 310

Search completed: August 20, 2003, 12:45:15
Job time : 1.73494 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 4.19277 Seconds

(Without alignments)
137.621 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIPP 6

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	359	2 A42462	acetoal catabolism
2	33	97.1	598	2 B40713	cylicin I - human
3	33	97.1	667	2 A40713	cylicin I - bovine
4	32	94.1	138	1 SXAD97	hexon-associated p
5	32	94.1	138	1 SXAD93	hexon-associated p
6	32	94.1	393	2 A55859	regulatory protein
7	31	91.2	306	2 H81036	riboflavin kinase/
8	31	91.2	318	2 A81982	FAD synthase NMA06
9	31	91.2	342	2 S23764	polygalacturonase-
10	31	91.2	613	2 H71803	penicillin-binding
11	31	91.2	615	2 D64714	cell division prot
12	31	91.2	671	2 S51599	Om(2D) protein - f
13	31	91.2	744	2 C84527	probable receptor-
14	31	91.2	777	2 C96745	hypothetical prote
15	31	91.2	981	2 T50851	receptor protein k
16	31	91.2	1002	2 T46033	receptor protein k
17	31	91.2	1196	2 T09356	brassinosteroid-in
18	30	88.2	146	2 C70920	hypothetical prote
19	30	88.2	210	2 C81216	conserved hypothet
20	30	88.2	210	2 G81793	hypothetical prote
21	30	88.2	244	2 T39507	probable ATP synth
22	30	88.2	293	2 T01258	hypothetical prote
23	30	88.2	343	2 A82641	UDP-galactose 4-ep
24	30	88.2	343	2 H97423	UDP-galactose 4-ep
25	30	88.2	354	2 C81053	alpha-1,2-N-acetyl
26	30	88.2	354	2 E81824	alpha-1,2-N-acetyl
27	30	88.2	472	2 AD1889	adenylate cyclase
28	30	88.2	604	2 T04351	viviparous-14 prot
29	30	88.2	932	2 T48489	receptor-like prot

30	30	88.2	1109	2 T18536	receptor-like prot
31	30	88.2	1123	2 D96756	receptor-like prot
32	30	88.2	2324	1 A29924	acetyl-CoA carboxy
33	30	88.2	2339	2 S41121	acetyl-CoA carboxy
34	30	88.2	2346	2 I38928	acetyl-CoA carboxy
35	30	88.2	4767	2 T31345	hypothetical prote
36	29	85.3	57	2 S10782	salivary protein p
37	29	85.3	75	2 A69010	hypothetical prote
38	29	85.3	79	1 PJH0SB	proline-rich pepi
39	29	85.3	86	2 D65181	hypothetical prote
40	29	85.3	99	2 B87374	transcription regu
41	29	85.3	100	2 F82453	conserved hypothet
42	29	85.3	121	2 G75604	hypothetical prote
43	29	85.3	133	2 S10038	hypothetical prote
44	29	85.3	134	2 JC5572	proline-rich prote
45	29	85.3	149	2 F91217	hypothetical prote

ALIGNMENTS

RESULT 1
A42462
acetoal catabolism protein ACOX - Alcaaligenes eutrophus (strain H16)
C:Species: Alcaaligenes eutrophus
C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C:Accession: A42462
R:Prilefert, H.; Hein, S.; Krueger, N.; Zeh, K.; Schmidt, B.; Steinbuechel, A.
J. Bacteriol. 173, 4056-4071, 1991
A:Title: Identification and molecular characterization of the Alcaaligenes eutrophus H
A:Reference number: A42462; MUID:91286190; PMID:2061286
A:Accession: A42462
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <PRI>
A:Cross-references: GB:M66060

Query Match 100.0%; Score 34; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 170 TGRIPP 175

RESULT 2

B40713
cylicin I - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 21-Jul-2000
R:Hess, H.; Held, H.; Franke, W.W.
J. Cell Biol. 122, 1043-1052, 1993
A:Title: Molecular characterization of mammalian cylicin, a basic protein of the sper
A:Reference number: A40713; MUID:93359502; PMID:8354692
A:Accession: B40713
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-598 <HE>
A:Cross-references: GB:Z22780; NID:9396104; PIDN:CAA80457.1; PID:9396105
C:Keywords: cytoskeleton

Query Match 97.1%; Score 33; DB 2; Length 598;
Best Local Similarity 83.3%; Pred. No. 45;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 544 TGRIPP 549

RESULT 3

A40713
 cyliclin I - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 05-Nov-1999
 C:Accession: A40713; S35913
 R:Hess, H.; Heid, H.; Franke, W.W.
 J. Cell Biol. 122, 1043-1052, 1993
 A:Title: Molecular characterization of mammalian cyliclin, a basic protein of the sperm
 A:Reference number: A40713; MUID:93359502; PMID:8354692
 A:Accession: A40713
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-667 <RES>
 A:Cross-references: GB:222779; NID:9396150; PIDN:CAA80456.1; PID:9396151
 C:Keywords: cytoskeleton

Query Match 97.1%; Score 33; DB 2; Length 667;
 Best Local Similarity 83.3%; Pred. No. 51;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||:|
 DB 613 TGRVPP 618

RESULT 4
 SXAD97
 hexon-associated protein - human adenovirus 7
 N:Alternate names: polypeptide IX
 C:Species: Mastadenovirus h7 (human adenovirus 7)
 C:Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 05-Jun-1998
 C:Accession: A03854
 R:DiJkema, R.; Maat, J.; Dekker, B.M.M.; van Ormondt, H.; Boyer, H.W.
 Gene 13, 373-385, 1981
 A:Title: The gene for polypeptide IX of human adenovirus type 7.
 A:Reference number: A91480; MUID:81261948; PMID:6266923
 A:Accession: A03854
 A:Molecule type: DNA
 A:Residues: 1-138 <DID>
 C:Superfamily: adenovirus hexon-associated protein (IX)
 C:Keywords: hexon-associated protein

Query Match 94.1%; Score 32; DB 1; Length 138;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||:|
 DB 17 TGRLLP 22

RESULT 5
 SXAD93
 hexon-associated protein (IX) - human adenovirus 3
 C:Species: Mastadenovirus h3 (human adenovirus 3)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
 C:Accession: B03854; A03854
 R:Engler, J.A.
 Gene 13, 387-394, 1981
 A:Title: The nucleotide sequence of the polypeptide IX gene of human adenovirus type 3.
 A:Reference number: A91481; MUID:81261949; PMID:7262560
 A:Accession: B03854
 A:Molecule type: DNA
 A:Residues: 1-138 <ENG>
 A:Cross-references: GB:J01962; NID:9209966; PIDN:AAA42510.1; PID:9209967
 C:Superfamily: adenovirus hexon-associated protein (IX)
 C:Keywords: hexon-associated protein

Query Match 94.1%; Score 32; DB 1; Length 138;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6

DB 17 TGRLLP 22
 |||:|

RESULT 6
 A55859
 regulatory protein nasr - Klebsiella pneumoniae
 C:Species: Klebsiella pneumoniae
 C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 04-Mar-2000
 C:Accession: A55859
 R:Goldman, B.S.; Lin, J.T.; Stewart, V.
 J. Bacteriol. 176, 5077-5085, 1994
 A:Title: Identification and structure of the nasr gene encoding a nitrate- and nitrite
 1
 A:Reference number: A55859; MUID:94327477; PMID:8051020
 A:Accession: A55859
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <GOL>
 A:Cross-references: GB:L27824; NID:9530886; PIDN:AAA25101.1; PID:9551871
 C:Genetics:
 A:Gene: nasr
 C:Superfamily: Klebsiella pneumoniae regulatory protein nasr

Query Match 94.1%; Score 32; DB 2; Length 393;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||:|
 DB 119 TGRLLP 124

RESULT 7
 H81036
 riboflavin kinase/FMN adenylyltransferase NMB1834 [Imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: H81036
 R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. et al.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: H81036
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-306 <TER>
 A:Cross-references: GB:AE002534; GB:AE002098; NID:97227095; PIDN:AAF42169.1; PID:9722
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1834
 C:Superfamily: conserved hypothetical protein H10963

Query Match 91.2%; Score 31; DB 2; Length 306;
 Best Local Similarity 83.3%; Pred. No. 59;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||:|
 DB 66 TGRMPP 71

RESULT 8
 A81982
 RNA synthase NMA0621 [similarity] - Neisseria meningitidis (strain 22491 serogroup A)
 N:Contains: FMN adenylyltransferase (EC 2.7.7.2); riboflavin kinase (EC 2.7.1.26)
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: A81982
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: A81982
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-318 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PIDN:CAB83911.1; PID:9737935
 A:Experimental source: serogroup A, strain 22491
 C:Genetics:
 A:Gene: rfbP, NMA0621
 C:Superfamily: conserved hypothetical protein H10963
 C:Keywords: nucleotidyltransferase; phosphotransferase

Query Match 91.2% Score 31; DB 2; Length 318;
 Best Local Similarity 83.3% Pred. No. 61;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGRIPP 6
 DB 78 TGRMP 83

RESULT 9

S23764
 polygalacturonase-inhibiting protein precursor - kidney bean
 C:Species: *Phaseolus vulgaris* (kidney bean)
 C:Date: 05-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
 C:Accession: S23764
 R:Tombart, P.; Desiderio, A.; Salvà, G.; Cervone, F.; Daroda, L.; de Lorenzo, G.; Bergma
 plant J. 2, 367-373, 1992
 A:Title: Cloning and characterization of the gene encoding the endopolygalacturonase-inh
 A:Reference number: S23764; MUID:93272053; PMID:1303801
 A:Accession: S23764
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-342 <TOU>
 A:Cross-references: EMBL:X64769; NID:921028; PIDN:CAA46016.1; PID:921029
 A>Note: It is uncertain whether Met-1 or Met-10 is the initiator
 C:Superfamily: polygalacturonase-inhibiting protein; leucine-rich alpha-2-glycoprotein F:274-297/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LNR>

Query Match 91.2% Score 31; DB 2; Length 342;
 Best Local Similarity 83.3% Pred. No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGRIPP 6
 DB 216 TGRIPP 221

RESULT 10

H71803
 penicillin-binding protein - *Helicobacter pylori* (strain J99)
 C:Species: *Helicobacter pylori*
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 17-Mar-2000
 C:Accession: H71803
 R:Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.;
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: H71803
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-613 <ARN>
 A:Cross-references: GB:AE001568; GB:AE001439; NID:94156083; PIDN:AAD07044.1; PID:9415609
 C:Genetics:
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: jhp1464
 C:Superfamily: penicillin-binding protein 3

Query Match 91.2% Score 31; DB 2; Length 613;
 Best Local Similarity 83.3% Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGRIPP 6
 DB 421 TGRIPP 426

RESULT 11

D64714
 cell division protein - *Helicobacter pylori* (strain 26695)
 C:Species: *Helicobacter pylori*
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 17-Mar-2000
 C:Accession: D64714
 R:Tomb, J.E.; White, O.; Karlvage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Feldman, J.M.; Fujii, C.; Bowman, C.; Watthey,
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpe, P.D.; Smith, H.O.; Fraser,
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: D64714
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-615 <TOM>
 A:Cross-references: GB:AE000653; GB:AE000511; NID:92314733; PIDN:AAD08596.1; PID:9231
 C:Superfamily: penicillin-binding protein 3

Query Match 91.2% Score 31; DB 2; Length 615;
 Best Local Similarity 83.3% Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGRIPP 6
 DB 421 TGRIPP 426

RESULT 12

S51599
 Om(2D) protein - fruit fly (*Drosophila ananassae*)
 C:Species: *Drosophila ananassae*
 C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
 C:Accession: S51599
 R:Toshida, K.; Jun, N.; Awasaki, T.; Tsuruya, Y.; Shaya, N.; Horii, S.H.
 Mol. Gen. Genet. 245, 577-587, 1994
 A:Title: Retrotransposon-induced ectopic expression of the Om(2D) gene causes the eye
 A:Reference number: S51599; MUID:95107256; PMID:7808408
 A:Accession: S51599
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-671 <YOS>
 A:Cross-references: EMBL:D26553; NID:9443769; PID:di006095; PID:9443770
 A:Gene: FlyBase:Dana/Om(2D)
 A:Cross-references: FlyBase:FBgn0010400

Query Match 91.2% Score 31; DB 2; Length 671;
 Best Local Similarity 83.3% Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGRIPP 6
 DB 431 TGRIPP 436

RESULT 13

C84527
 Probable receptor-like protein kinase [Imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: C84527

R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: C84527
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-744 <STO>
 A:Cross-references: GB:AE002093; NID:94662631; PIDN:AAD26903.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g15300
 A:Map position: 2

Query Match 91.2%; Score 31; DB 2; Length 744;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 ||:||||
 Db 161 TGRIP 166

RESULT 14

C96745
 hypothetical protein T9N14.3 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96745
 R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96745
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-977 <STO>
 A:Cross-references: GB:AE005173; NID:910645382; PIDN:AAG21501.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T9N14.3
 A:Map position: 1

Query Match 91.2%; Score 31; DB 2; Length 977;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 ||:||||
 Db 255 TGRIP 260

RESULT 15

T50851
 receptor protein kinase homolog [imported] - soybean
 C:Species: Glycine max (soybean)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
 C:Accession: T50851
 R:Yamamoto, E.; Karakaya, H.C.; Knap, H.T.
 Biochim. Biophys. Acta 1491, 333-340, 2000
 A>Title: Molecular characterization of two soybean homologs of Arabidopsis thaliana CLAV
 A:Reference number: Z25262
 A:Accession: T50851
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-981 <YAM>

A:Cross-references: EMBL:AF197946; PIDN:AAF59905.1
 C:Genetics:
 A:Gene: CLV1A
 C:Superfamily: protein kinase xaz1; leucine-rich alpha-2-glycoprotein repeat homology

Query Match 91.2%; Score 31; DB 2; Length 981;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 ||:||||
 Db 471 TGRIP 476

Search completed: August 20, 2003, 12:42:28
 Job time : 6.19277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 2.27711 Seconds

(without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIP 6

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	359	1 ACOX_ALCEU	P27748 alcaligenes
2	34	100.0	359	1 GP18_HUMAN	O61344 mus sapien
3	34	100.0	659	1 SYT_THERH	P56881 thermus the
4	33	97.1	598	1 CYL1_HUMAN	P35663 homo sapien
5	33	97.1	667	1 CYL1_BOVIN	P35662 bos taurus
6	32	94.1	138	1 HEX9_ADE07	P33283 human adeno
7	32	94.1	372	1 GPDA_CUPLA	P52425 cuphea lanc
8	31	91.2	342	1 PG12_PHAVU	P58822 phaseolus v
9	31	91.2	342	1 PG12_PHAVU	P58823 phaseolus v
10	31	91.2	342	1 BR11_ARATH	O22476 arabidopsis
11	31	88.2	333	1 P5P2_HUMAN	O99933 mus musculu
12	30	88.2	387	1 T122_MOUSE	O64709 mus musculu
13	30	88.2	1135	1 RBL2_MOUSE	P11029 gallus gall
14	30	88.2	2346	1 COAC_CHICK	O91533 dos taurus
15	30	88.2	2346	1 COAL_BOVIN	O13085 homo sapien
16	30	88.2	2346	1 COAL_BOVIN	O28559 ovis aries
17	30	88.2	2346	1 COAL_BOVIN	P02814 homo sapien
18	30	88.2	2346	1 COAL_BOVIN	O99954 homo sapien
19	29	85.3	79	1 PRL3_HUMAN	P56259 escherichia
20	29	85.3	134	1 PRL3_HUMAN	P59151 woodwardia
21	29	85.3	149	1 YIFN_ECOLI	O99137 thermotoga
22	29	85.3	185	1 R44_MOORA	O07276 macaca fasc
23	29	85.3	195	1 ENGB_THEMA	O57891 methanococc
24	29	85.3	257	1 KLR1_MACFA	P40794 drosophila
25	29	85.3	283	1 Y449_METJA	O34136 drosophila
26	29	85.3	303	1 BNK_DROME	O99136 chlamydia m
27	29	85.3	307	1 DNAL_DEIPR	P25438 xanthomonas
28	29	85.3	317	1 HEMZ_CHIMU	P78814 schizosacch
29	29	85.3	346	1 YR61_XANCV	O00874 arabidopsis
30	29	85.3	365	1 SR2_SCHPO	O24106 drosophila
31	29	85.3	372	1 D100_ARATH	
32	29	85.3	410	1 HID_DROME	
33	29	85.3	410	1 HID_DROME	

34	29	85.3	413	1 CYL1_DROME	O61734 drosophila
35	29	85.3	420	1 CT13_HUMAN	O9h6p5 homo sapien
36	29	85.3	420	1 CT13_MOUSE	O81191 mus musculu
37	29	85.3	450	1 YC14_METRA	O8r623 methanopyru
38	29	85.3	492	1 YGFH_ECOLI	P52043 escherichia
39	29	85.3	624	1 STS_MOUSE	P50427 mus musculu
40	29	85.3	628	1 V7OK_TYMW	P10357 turnip yell
41	29	85.3	628	1 V7OK_TYMW	P28478 turnip yell
42	29	85.3	693	1 CAUP_DROME	P54269 drosophila
43	29	85.3	700	1 YDHY_ECOLI	P76192 escherichia
44	29	85.3	705	1 MMUC_STROO	O88022 streptomyce
45	29	85.3	738	1 PARC_RICCN	O92j10 rickettsia

ALIGNMENTS

RESULT 1

ACOX_ALCEU STANDARD; PRT; 359 AA.

AC P27748:

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Acetoin catabolism protein X.

GN ACOX.

OS Alcaligenes eutrophus (Ralstonia eutropha).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OX Burkholderiaceae; Ralstonia.

NCBI_TaxID=510;

RP SEQUENCE FROM N.A.

RC STRAIN-H16 / DSM 428 / ATCC 17699;

RX MEDLINE=91286190; PubMed=2061286;

RA Pfeleert H., Hein S., Krueger N., Zeh K., Schmidt B.,

RT Steinbuechel A.;

RT "Identification and molecular characterization of the Alcaligenes

RT eutrophus H16 aco operon genes involved in acetoin catabolism.";

RL J. Bacteriol. 173:4056-4071(1991).

CC - FUNCTION: ESSENTIAL FOR ACETOIN CATABOLISM.

CC - SUBCELLULAR LOCATION: MEMBRANE (POTENTIAL).

CC - This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL: M6060; AAA1947.1; -

KW Acetoin catabolism; Membrane.

SQ SEQUENCE 359 AA; 37934 MW; 6825322F5BHC50E5 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 359;

Best Local Similarity 100.0%; Pred. No. 6.4;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6

DB 170 TGRIP 175

RESULT 2

GP18_HUMAN STANDARD; PRT; 395 AA.

AC O92643; 014822;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE GPI-anchor transamidase (EC 3.-.-.-) (GPI transamidase)

(Phosphatidylinositol-glycan biosynthesis, class K protein) (Ptc-K)

DE (hsp18).
GN PIGK OR GP18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97133287; PubMed=8978684;
RA Benghezal M., Beauchour A., Rusconi S., Aebi M., Conzelmann A.;
RT "Yeast Gpi1p is essential for GPI anchor attachment onto proteins.";
RL EMO J. 15:6575-6583(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98024174; PubMed=9356492;
RA Yu J., Nagatajan S., Knež J.J., Udenfriend S., Chen R., Medof M.E.;
RT "The affected gene underlying the class K glycosylphosphatidylinositol
(R) (GPI) surface protein defect codes for the GPI transamidase.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:12580-12585(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TRISQUE-LIVER;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udsal T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollaly S.J.,
RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield A.S.N., Krzywinski M.I., Skalska U., Smolins D.E.,
RA Schnerch A., Schein J.E., Jones S.U.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Mediates GPI anchoring in the endoplasmic reticulum, by
replacing a protein's C-terminal GPI attachment signal peptide
with a pre-assembled GPI. During this transamidation reaction, the
GPI transamidase forms a covalent intermediate with the substrate
protein.
CC -1- PATHWAY: GPI-anchor biosynthesis.
CC -1- SUBUNIT: Associates with PIGS and PIGT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y07596; CAA68871.1; -
CC DR EMBL: AF022913; AAB81597.1; -
CC DR EMBL: BC020737; AAH20737.1; -
CC DR Genew: HGNC:8965; PIGK.
CC DR MIM: 605087; -
CC DR MEROPS: C13.005; -
CC DR InterPro: IPR001096; Legumain.
CC DR Pfam: PF01650; Peptidase_C13; 1.
CC DR PRINTS: PR00776; HEMOGLOBINASE.
CC DR Hydrolase: Thiol protease; GPI-anchor biosynthesis.
CC ACT_SITE 164 164
CC FT ACT_SITE 206 206 POTENTIAL.
CC FT CONFLICT 1 4 MAY1 -> SLH8A (IN REF. 1).
CC

SO SEQUENCE 395 AA; 45251 MW; AF706DDAD13EEF2 CRC64;
Query Match 100.0%; Score 34; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGRIP 6
DB 136 TGRIP 141
RESULT 3
SYT_THETH STANDARD; PRT; 659 AA.
AC P56881.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine--tRNA ligase)
DE (ThRS).
GN THRS.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Thermococci; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HB8 / ATCC 27634;
RA CURA V., Moras D., Kern D.;
RT "Sequence analysis and modular organization of threonyl-tRNA
synthetase from Thermus thermophilus and its interrelation with
threonyl-tRNA synthetases of other origins.";
RL Eur. J. Biochem. 267:379-393(2000).
RN [2]
RP SEQUENCE OF 1-45, AND CHARACTERIZATION.
RX STRAIN=HB8 / ATCC 27634;
RA MEDLINE=9430498; PubMed=8031907;
RA Zheltosova J., Melnikova E., Garber M., Reinbolt J., Kern D.,
RA Ehrsman C., Ehrsman B.;
RT "Threonyl-tRNA synthetase from Thermus thermophilus: purification and
some structural and kinetic properties.";
RL Biochimie 76:71-77(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-threonine + tRNA(Thr) = AMP +
diphosphate + L-threonyl-tRNA(Thr).
CC -1- COFACTOR: Binds 1 zinc ion per subunit.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DOMAIN: THE C-TERMINAL DOMAIN RECOGNIZES THE ANTICODON BASES.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ250953; CAB65483.1; -
CC DR HSSP: P00955; IEVL.
CC DR HAMAP: MF_00184; -; 1.
CC DR InterPro: IPR004154; HGTP_anticodon.
CC DR InterPro: IPR004095; TGS_dom.
CC DR InterPro: IPR002314; tRNA-synt_2b.
CC DR InterPro: IPR002320; tRNA-synt_thr.
CC DR InterPro: IPR006195; tRNA_ligase_II.
CC DR Pfam: PF03129; HGTP_anticodon; 1.
CC DR Pfam: PF02824; TGS; 1.
CC DR Pfam: PF00587; tRNA-synt_2b; 1.
CC DR PRINTS: PR01047; TRNASYNTHTHR.
CC DR TIGRFAMS: TIGR00418; THRS; 1.
CC PROSITE: PSS0862; AA_TRNA_LIGASE_II; 1.
CC

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-Binding;
 KM Metal-binding; Zinc.
 FT DOMAIN 234 548 CATALYTIC.
 FT METAL 338 341 POLY-GLU.
 FT METAL 349 349 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 400 400 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 529 529 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CONFLICT 42 42 E -> T (IN REF. 2).
 SQ SEQUENCE 659 AA; 75537 MW; 1993A1BEF151FF4 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 659;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 198 TGRIPP 203

RESULT 4
 CYL_HUMAN STANDARD; PRT; 598 AA.
 ID CYL_HUMAN
 AC P35663;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cylclin I (Multiple-band polypeptide I) (Fragment).
 GN Cylcl I OR CYL OR CYL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93359502; PubMed=8354692;
 RA Hess H., Heid H., Franke W.W.;
 RT "Molecular characterization of mammalian cylclin, a basic protein of
 the sperm head cytoskeleton.";
 RT J. Cell Biol. 123:1043-1052(1993).
 CC -1- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY
 BE INVOLVED IN SPERMATID DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.
 CC -1- TISSUE SPECIFICITY: Testis.
 CC -1- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 222780; CAA80457.1; -;
 DR PIR; BA0713; B40713.
 DR Gene; HGNC:2582; CYLC1.
 DR MIM; 603121;
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0005198; F:structural molecule activity; NAS.
 DR GO; GO:0007283; P:spermatogenesis; NAS.
 KW Cytoskeleton; Structural protein; Repeat; Sperm; Spermatogenesis.
 FT NON_TER 1 1
 FT DOMAIN 479 500 8 APPROXIMATE TANDEM REPEATS.
 FT REPEAT 225 252 1.
 FT REPEAT 253 289 2.
 FT REPEAT 290 326 3.
 FT REPEAT 327 364 4.
 FT REPEAT 365 400 5.
 FT REPEAT 401 438 6.
 FT REPEAT 439 478 7.
 FT REPEAT 479 500 8.
 FT DOMAIN 548 598 PRO-RICH.

SQ SEQUENCE 598 AA; 68034 MW; 062BA2E2D2AB1F7 CRC64;
 Query Match 97.1%; Score 33; DB 1; Length 598;
 Best Local Similarity 83.3%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 544 TGRIPP 549

RESULT 5
 CYL_BOVIN STANDARD; PRT; 667 AA.
 ID CYL_BOVIN
 AC P35662;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Cylclin I (Multiple-band polypeptide I).
 GN Cylcl I OR CYL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=93359502; PubMed=8354692;
 RA Hess H., Heid H., Franke W.W.;
 RT "Molecular characterization of mammalian cylclin, a basic protein of
 the sperm head cytoskeleton.";
 RT J. Cell Biol. 123:1043-1052(1993).
 CC -1- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY
 BE INVOLVED IN SPERMATID DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.
 CC -1- TISSUE SPECIFICITY: Testis.
 CC -1- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 222779; CAA80456.1; -;
 DR PIR; A40713; A40713.
 DR CYTOSKELETON; Structural protein; Repeat; Sperm; Spermatogenesis.
 KW CYTOSKELETON; Structural protein; Repeat; Sperm; Spermatogenesis.
 FT DOMAIN 287 569 9 APPROXIMATE TANDEM REPEATS.
 FT REPEAT 287 305 1.
 FT REPEAT 306 337 2.
 FT REPEAT 338 368 3.
 FT REPEAT 369 405 4.
 FT REPEAT 406 442 5.
 FT REPEAT 443 475 6.
 FT REPEAT 476 516 7.
 FT REPEAT 517 547 8.
 FT REPEAT 548 569 9.
 FT DOMAIN 617 667 PRO-RICH.
 SQ SEQUENCE 667 AA; 74817 MW; CBF66EA62243D91 CRC64;

Query Match 97.1%; Score 33; DB 1; Length 667;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 613 TGRIPP 618

RESULT 6

```

HEX9_ADE07
ID HEX9_ADE07 STANDARD; PRT; 138 AA.
AC P03283;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hexon-associated protein (Protein IX).
GN PIX.
OS Human adenovirus type 7, and
OS Human adenovirus type 3.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519, 45659;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human adenovirus type 7; STRAIN=Gomen;
RX MEDLINE=81261948; PubMed=6266923;
RA Dijkema R., Maat J., Dekker B.M.M., van Ormondt H., Boyer H.W.;
RT "The gene for polypeptide IX of human adenovirus type 7."
RL Gene 13:375-385(1981).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human adenovirus type 3;
RX MEDLINE=81261949; PubMed=7262560;
RA Engler J.A.;
RT "The nucleotide sequence of the polypeptide IX gene of human
adeno virus type 3."
RL Gene 13:387-394(1981).
CC -1- FUNCTION: THIS PROTEIN IS A STRUCTURAL COMPONENT OF THE VIRION. IT
MAY HAVE AN ADDITIONAL ROLE DURING ADENOVIRUS MULTIPLICATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

```

```

RT RT Cloning of a cDNA coding for a glycerol-3-phosphate dehydrogenase
RT from Cuphea lanceolata.
RT (In) Kader J.-C., Mazliak P. (eds.);
RL Plant lipid metabolism, pp.534-536, Kluwer Academic Publishers,
RL Dordrecht (1995).
CC -1- CATALYTIC ACTIVITY: Sn glycerol 3-phosphate + NAD(+) -> glycerone
phosphate + NADH.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE
DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

```

RT confers a new recognition capability.";
 RL EMO J. 18:2352-2363(1999).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=97449843; PubMed=9304859;
 RA Desiderio A., Aracri B., Leckie F., Mattel B., Salvi G., Tigelaar H.,
 RA Van Roekel J.S., Baulcombe D.C., Melchers L.S., De Lorenzo G.,
 RA Cervone F.;
 RT "Polygalacturonase-inhibiting proteins (PGIPs) with different
 RT specificities are expressed in *Phaseolus vulgaris*.";
 RL Mol. Plant Microbe Interact. 10:852-860(1997).
 CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an
 CC important factor for plant resistance to phytopathogenic fungi.
 CC Substrate preference is polygalacturonase (PG) from *A.niger* >> PG
 CC of *F.oxysporum*, *A.solani* or *B.cinerea*. Not active on PG from
 CC *F.moniliforme*.
 CC -1- SUBCELLULAR LOCATION: Cell-wall associated.
 CC -1- MISCELLANEOUS: Mutation of Lys-253 confers the ability to inhibit
 CC the *F.moniliforme* PG.
 CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITING PROTEIN
 CC FAMILY.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-10 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X64769; CAA016.1; -;
 DR EMBL; A23205; CAA01664.1; -;
 DR PIR; S23764; S23764.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_plant.
 DR Pfam; PF00560; LRR; 4.
 KW Signal; Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 29
 FT CHAIN 30 342
 FT REPEAT 78 104 LRR 1.
 FT REPEAT 129 152 LRR 2.
 FT REPEAT 153 177 LRR 3.
 FT REPEAT 179 200 LRR 4.
 FT REPEAT 201 225 LRR 5.
 FT REPEAT 272 296 LRR 6.
 FT REPEAT 298 321 LRR 7.
 FT DISULFID 32 62 BY SIMILARITY.
 FT DISULFID 63 72 BY SIMILARITY.
 FT DISULFID 310 332 BY SIMILARITY.
 FT DISULFID 334 341 BY SIMILARITY.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MUTAGEN 253 253 K->Q: BROWDER SPECTRUM OF ACTION.
 SQ SEQUENCE 342 AA; 37101 MW; 950F94E0D2A39598 CRC64;
 Query Match 91.28; Score 31; DB 1; Length 342;
 Best Local Similarity 83.38; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Polygalacturonase inhibitor 2 precursor (polygalacturonase-inhibiting
 DE protein) (PGIP-2).
 GN PGIP2.
 OS Phaseolus vulgaris (kidney bean) (French bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxId=3885;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.
 RP STRAIN=cv. Pinto; TISSUE=Hypocotyl;
 RX MEDLINE=99246261; PubMed=10228150;
 RA Leckie F., Mattel B., Capoditosa C., Hemmings A., Nuss L., Aracri B.,
 RA De Lorenzo G., Cervone F.;
 RT "The specificity of polygalacturonase-inhibiting protein (PGIP): a
 RT single amino acid substitution in the solvent-exposed
 RT beta-strand/beta-turn region of the leucine-rich repeats (LRRs)
 RT confers a new recognition capability.";
 RL EMO J. 18:2352-2363(1999).
 RN [2]
 RP STRUCTURE OF CARBOHYDRATES ON ASN-64 AND ASN-141, AND DISULFIDE BONDS.
 RX MEDLINE=20582621; PubMed=11148052;
 RA Mattel B., Bernaldi M.S., Federici L., Roepstorff P., Cervone F.,
 RA Boffi A.;
 RT "Secondary structure and posttranslational modifications of the
 RT leucine-rich repeat protein PGIP (polygalacturonase-inhibiting
 RT protein) from *Phaseolus vulgaris*.";
 RL Biochemistry 40:569-576(2001).
 CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an
 CC important factor for plant resistance to phytopathogenic fungi.
 CC Inhibits all polygalacturonases (PG) tested, with the exception of
 CC PG from *F.oxysporum* which was only inhibited at 60%.
 CC -1- SUBCELLULAR LOCATION: Cell-wall associated.
 CC -1- MISCELLANEOUS: Asn-303 is not glycosylated.
 CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITING PROTEIN
 CC FAMILY.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-10 is the initiator.
 CC Glycosylated; P58822; -;
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_plant.
 DR Pfam; PF00560; LRR; 5.
 KW Signal; Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 29
 FT CHAIN 30 342
 FT REPEAT 78 104 LRR 1.
 FT REPEAT 129 152 LRR 2.
 FT REPEAT 153 177 LRR 3.
 FT REPEAT 179 200 LRR 4.
 FT REPEAT 201 225 LRR 5.
 FT REPEAT 272 296 LRR 6.
 FT REPEAT 298 321 LRR 7.
 FT DISULFID 32 62 PROBABLE.
 FT DISULFID 63 72 PROBABLE.
 FT DISULFID 310 332
 FT DISULFID 334 341
 FT CARBOHYD 64 64
 FT CARBOHYD 141 141
 FT CARBOHYD 303 303
 FT MUTAGEN 181 181
 FT MUTAGEN 207 207
 FT MUTAGEN 253 253
 FT MUTAGEN 300 300
 FT MUTAGEN 320 320
 FT MUTAGEN 326 326
 FT MUTAGEN 340 340
 N-LINKED (GLCNAC. . .) (COMPLEX).
 N-LINKED (GLCNAC. . .) (COMPLEX).
 L->H: NO EFFECT.
 V->G: NO EFFECT; LOSS OF ACTIVITY WHEN
 ASSOCIATED WITH K-253; NO EFFECT WHEN
 ASSOCIATED WITH S-326.
 S->A: NO EFFECT.
 O->K: 70% DECREASE OF ACTIVITY; LOSS OF
 ACTIVITY WHEN ASSOCIATED WITH G-181 OR S-
 326.
 H->Q: NO EFFECT.
 Q->K: NO EFFECT.
 A->S: NO EFFECT; NO EFFECT WHEN
 ASSOCIATED WITH G-181; LOSS OF ACTIVITY
 WHEN ASSOCIATED WITH K-253.
 A->S: NO EFFECT.

SQ SEQUENCE 342 AA; 37067 MW; 8B93E4AEBA46477E CRC64;

Query Match 91.2%; Score 31; DB 1; Length 342;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 ||:||||
 DB 216 TGRIPP 221

RESULT 10
 PG13_PHAVU STANDARD; PRT; 342 AA.
 AC P58823;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Polygalacturonase inhibitor 3 precursor (Polygalacturonase-inhibiting protein) (PGIP-2) (PGIP-3).
 GN Pgip3.
 OS Phaseolus vulgaris (kidney bean) (French bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxID=3885;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. Pinto; TISSUE=HYPOCOTYL;
 RX MEDLINE=93372053; PubMed=1303801;
 RA Toubart P., Desiderio A., Salvi G., Ceryone F., Daroda L., de Lorenzo G., Bergmann C., Darvill A.G., Alberheim P.;
 RT "Cloning and characterization of the gene encoding the endopolygalacturonase-inhibiting protein (PGIP) of Phaseolus vulgaris L.";
 RT Plant J. 2:367-373(1992).
 RL
 CC -1- FUNCTION: Inhibitor of fungal polygalacturonase. It is an important factor for plant resistance to phytopathogenic fungi.
 CC -1- SUBCELLULAR LOCATION: Cell-wall associated.
 CC -1- TISSUE SPECIFICITY: Found in suspension-cultured cells and to a lesser extent in hypocotyls, leaves and flowers.
 CC -1- SIMILARITY: BELONGS TO THE POLYGALACTURONASE-INHIBITING PROTEIN FAMILY.
 CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-10 is the initiator.
 DR Interpro: IPR001611; LRR.
 DR Interpro: IPR007090; LRR_plant.
 DR Pfam: PF00560; LRR; 5.
 KW Signal; Glycoprotein; Cell wall; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 29
 FT CHAIN 30 342 POLYGALACTURONASE INHIBITOR 3.
 FT REPEAT 78 104 LRR 1.
 FT REPEAT 129 152 LRR 2.
 FT REPEAT 153 177 LRR 3.
 FT REPEAT 179 200 LRR 4.
 FT REPEAT 201 225 LRR 5.
 FT REPEAT 272 296 LRR 6.
 FT REPEAT 298 321 LRR 7.
 FT DISULFID 32 62 BY SIMILARITY.
 FT DISULFID 63 72 BY SIMILARITY.
 FT DISULFID 310 332 BY SIMILARITY.
 FT DISULFID 334 341 BY SIMILARITY.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 342 AA; 37251 MW; 980AEFID3A28F27D CRC64;

Query Match 91.2%; Score 31; DB 1; Length 342;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 ||:||||

DB 216 TGRIPP 221

RESULT 11
 BR11_ARATH STANDARD; PRT; 1196 AA.
 ID BR11_ARATH
 AC 022476;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE BRASSINOSTEROID INSENSITIVE 1 precursor (EC 2.7.1.37) (AtBRI1)
 DE (Brassinosteroid LRR receptor kinase).
 GN BRI1 OR Atg39400 OR F23K16.30.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE=SPECIFICITY, AND MUTANTS BRI1-101;
 RP BRI1-104; BRI1-113 AND BRI1-115.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=97442355; PubMed=9298904;
 RA Li J., Chory J.;
 RT "A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction.";
 RT Cell 90:929-938(1997).
 RL
 RN [2]
 RP SEQUENCE FROM N.A., FUNCTION, AND MUTANTS BRI1-5/DWFE2-W41;
 RP BRI1-6/BRI1-119/DWFE2-399; BRI1-7/DWFE2-WM3-2; BRI1-8/DWFE2-WM6-2 AND BRI1-9/DWFE2-WM19.
 RP STRAIN=cv. Wassilewskija-2, and cv. En-2;
 RC MEDLINE=20027415; PubMed=10557222;
 RX Noguchi T., Fujioka S., Choe S., Takatsuto S., Yoshida S., Yuan H., Feldmann K.A., Tax F.E.;
 RA "Brassinosteroid-insensitive dwarf mutants of Arabidopsis accumulate brassinosteroids.";
 RT Plant Physiol. 121:743-752(1999).
 RL
 RN [3]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MUTANTS BRI1-1; BRI1-108; BRI1-117 AND BRI1-102.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20398322; PubMed=10938344;
 RA Friedlrichsen D.M., Joazeiro C.A.P., Li J., Hunter T., Chory J.;
 RT "BRASSINOSTEROID-INSENSITIVE-1 is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase.";
 RT Plant Physiol. 123:1247-1256(2000).
 RL
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Ertlan K.-D., Terry N., Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M., Kreis M., Delseny M., Pulgomech P., Watson M., Schmidheini T., Reichert B., Portetlelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hohnsels J., Zimmermann W., Wedler H., Ridley P., Landham S.-A., McCullagh B., Bilham L., Robben J., Van der Schuren J., Grymoprez B., Chuang Y.-J., Vandenbusche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenegeger T., Botne G., Ramsperger U., Hilbert H., Braun M., Holzner E., Brandt A., Peters S., van Staveren M., Dirse M., Moollman P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P., Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Brey-Allen S., Clark L., Doggett J., Hall S., Kay W., Lennard N., McIay K., Mayes R., Pettelt A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer K., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A., Neumann S., Argitirov A., Vitale D., Liguri R., Piravandi E.,

RA Massenet O., Quigley F., Clabaud G., Muehlen A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharry A., Aubourg S.,
RA Chebrolu F., Cooke R., Berger C., Montfort A., Casachuberta E.,
RA Gibbons T., Weber N., Vandendol M., Barques M., Tercel J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Balke C.,
RA Frishman D., Haas D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parrell L., Dedha N., Gnoj J., Schurz K., Huang E., Spiegel L.,
RA Storek M., Murray J., Sheet P., Cordes M., Abu-Theladeh J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speith J., Ryan E., Andrews S., Gelsel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Antonou B., Zidenic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martensen R., Mccomb W.R.;
RA "Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.";
RA Nature 402:769-777(1999).
RL [5]
RN PHOSPHORYLATION.
RX MEDLINE-20484044; PubMed-11027724;
RA Oh M.-H., Ray W.K., Huber S.C., Asara J.M., Gage D.A., Clouse S.D.;
RT "Recombinant BRASSINOSTEROID INSENSITIVE 1 receptor-like kinase
RT autophosphorylates on serine and threonine residues and
RT phosphorylates a conserved peptide motif *in vitro*.";
RL Plant Physiol. 124:751-766(2000).
RN [6]
RP STEROID BINDING.
RX MEDLINE-20336852; PubMed-10875920;
RA He Z., Wang Z.-Y., Li J., Zhu Q., Lamb C., Ronald P., Chory J.;
RT "Perception of brassinosteroids by the extracellular domain of the
RT receptor kinase BRI1.";
RL Science 288:2360-2363(2000).
RN [7]
RP SUBCELLULAR LOCATION, STEROID BINDING, AND AUTOPHOSPHORYLATION.
RX MEDLINE-21168182; PubMed-11268216;
RA Wang Z.-Y., Seto H., Fujioaka S., Yoshida S., Chory J.;
RT "BRI1 is a critical component of a plasma-membrane receptor for plant
RT steroids.";
RL Nature 410:380-383(2001).
RN [8]
RP SUBCELLULAR LOCATION, PHOSPHORYLATION, AND INTERACTION WITH BAK1.
RX MEDLINE-22145926; PubMed-12150928;
RA Nam K.H., Li J.;
RT "BRI1/BAK1, a receptor kinase pair mediating brassinosteroid
RT signaling.";
RL Cell 110:203-212(2002).
RN [9]
RP PHOSPHORYLATION, AND INTERACTION WITH BAK1.
RX MEDLINE-22145927; PubMed-12150929;
RA Li J., Wen J., Lease K.A., Duke J.T., Tax F.E., Walker J.C.;
RT "BAK1, an Arabidopsis LRR receptor-like protein kinase, interacts with
RT BRI1 and modulates brassinosteroid signaling.";
RL Cell 110:213-222(2002).
RN [10]
RP FUNCTION: Receptor with a serine/threonine-protein kinase
CC activity. Regulates, in response to brassinosteroid binding, a
CC signaling cascade involved in plant development, including
CC elongation of light- and stress-regulated genes, promotion of cell
CC elongation, normal leaf and chloroplast senescence, and flowering.
CC 2,3,22,23-O-tetramethylbrassinolide or ecdysone. May be involved
CC in a feedback regulation of brassinosteroid biosynthesis.
CC Phosphorylates BRI1-associated receptor kinase 1 (BAK1).
CC -1- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphorylated
CC -1- ENZYME REGULATION: Activated by Ser and Thr phosphorylation.
CC -1- SUBUNIT: Heterodimer with BAK1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein; plasma membrane.

CC -1- TISSUE SPECIFICITY: Expressed ubiquitously.
CC -1- DEVELOPMENTAL STAGE: Expressed constitutively in either dark- or
CC light-grown seedlings.
CC -1- DOMAIN: Contains one leucine-zipper motif and two pairs of
CC conservatively spaced Cys (Cys pair 1 and 2) involved in forming
CC heterodimers.
CC -1- PTM: Phosphorylated on at least 12 sites, with a preference for
CC Ser residues.
CC -1- MISCELLANEOUS: A 70 amino acid island between the 20th and the
CC 21th LRR is essential for the binding of brassinosteroids.
CC -1- MISCELLANEOUS: Binding of brassinosteroid induces intramolecular
CC autophosphorylation of BRI1. Interaction with BAK1 activates both
CC receptor kinases and the full activation of either receptor kinase
CC requires transphosphorylation by their partners. Optimum *in vitro*
CC phosphorylation of the substrate requires Arg or Lys residues at
CC P-3, P-4, and P+5 (relative to the phosphorylated amino acid at
CC P-0).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: Contains 25 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF017056; AAC49810.1; -;
CC DR EMBL: AL078620; CAB4675.1; -;
CC DR EMBL: AL161595; CAB80603.1; -;
CC DR PIR: T09356; T09356.
CC DR InterPro: IPR001611; LRR.
CC DR InterPro: IPR007090; LRR_plant.
CC DR InterPro: IPR007719; Prot_kinase.
CC DR InterPro: IPR002290; Ser_thr_kinase.
CC DR InterPro: IPR001245; Tyr_pkinase.
CC DR Pfam: PF00560; LRR_17.
CC DR Pfam: PF00069; pkinase_1.
CC DR PRINTS: PR00019; LEURICRPT.
CC DR PRINTS: PR00109; TYRKINASE.
CC DR PRODOM: PD000001; Prot_kinase; 1.
CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC DR Transferrase: Kinase; Serine/threonine-protein kinase; Receptor;
CC KW Steroid-binding; ATP-binding; Repeat; Signal; Transmembrane;
CC KW Leucine-rich repeat; Glycoprotein; Phosphorylation.
CC FT SIGNAL 1 23
CC FT CHAIN 24 1196 BRASSINOSTEROID INSENSITIVE 1.
CC FT TRANSMEM 793 813 POTENTIAL.
CC FT REPEAT 98 121 LRR 1.
CC FT REPEAT 122 146 LRR 2.
CC FT REPEAT 148 169 LRR 3.
CC FT REPEAT 172 197 LRR 4.
CC FT REPEAT 199 221 LRR 5.
CC FT REPEAT 222 244 LRR 6.
CC FT REPEAT 245 268 LRR 7.
CC FT REPEAT 269 290 LRR 8.
CC -----
CC Query Match 91.2%; Score 31; DB 1; Length 1196;
CC Best Local Similarity 83.3%; Pred. No. 1e+02;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC OY 1 TGRIPP 6
CC DB 429 TGRIPP 434
CC -----
CC RESULT 12
CC PSP2_HUMAN
CC ID PSP2_HUMAN STANDARD; PRT; 333 AA.
CC AC Q9H939;

ID	REL2_MOUSE	STANDARD;	PRT;	1135 AA.
AC	064700;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DE	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Retinoblastoma-like protein 2 (130 kDa retinoblastoma-associated protein) (PRB2) (P130) (RBR-2).			
GN	RBL2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_TaxID=10090;				
RA	Chen G., Guy C.T., Chen H.W., Hu N., Lee E.Y.H.P., Lee W.H.;			
RA	"Molecular cloning and developmental expression of mouse p130, a member of the retinoblastoma gene family.";			
RT	J. Biol. Chem. 271:9567-9572(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96199216; PubMed=8621630;			
RA	Chen G., Guy C.T., Chen H.W., Hu N., Lee E.Y.H.P., Lee W.H.;			
RA	"Molecular cloning and developmental expression of mouse p130, a member of the retinoblastoma gene family.";			
RT	J. Biol. Chem. 271:9567-9572(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96192345; PubMed=9019172;			
RA	Fertile P., Baldi A., de Luca A., Virgilio L., Pisano M.M.,			
RA	Giordano A.;			
RT	"Molecular cloning, expression, and developmental characterization of the murine retinoblastoma-related gene Rb2/p130.";			
RT	Cell Growth Differ. 6:1659-1664(1995).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96203998; PubMed=8622859;			
RA	Lecuter J.E., Whyte P.F.M., Rudnicki M.A.;			
RA	"Cloning and expression of the Rb-related mouse p130 mRNA.";			
RT	Oncogene 12:1433-1440(1996).			
CC	-1- FUNCTION: MAY HAVE A FUNCTION IN CELL CYCLE REGULATION. BINDS TO AND MAY BE INVOLVED IN THE TRANSFORMING CAPACITY OF THE ADENOVIRUS E2A PROTEIN. MAY ACT AS A TUMOR SUPPRESSOR. POTENT INHIBITOR OF E2F-MEDIATED TRANS-ACTIVATION. ASSOCIATES PREFERENTIALLY WITH E2F5. BINDS TO CYCLIN A AND E (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- PTM: Phosphorylation on Ser-669 in G1 leads to its ubiquitin-dependent proteolysis (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE RETINOBLASTOMA PROTEIN (RB) FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U36799; AAB48991.1; -;			
DR	EMBL; U50850; AAC52598.1; -;			
DR	EMBL; U47333; AAC52555.1; -;			
DR	HSSP; P06400; 1GUX.			
DR	TRANSFAC; T02972.			
DR	MGP; MGI:105085; Rb12.			
DR	InterPro; IPR006670; Cyclin.			
DR	InterPro; IPR002720; RB_A.			
DR	InterPro; IPR002719; RB_B.			
DR	Pfam; PF01858; RB_A; 1.			
DR	Pfam; PF01857; RB_B; 1.			
DR	SMART; SM00385; CYCLIN; 2.			
DR	Transcription regulation; DNA-binding; Nuclear protein; Cell cycle; Phosphorylation; Anti-oncogene.			
KW	POCKET (BINDS E2F).			
FT	DOMAIN 414 1021			
FT	DOMAIN 414 613			
FT	DOMAIN 614 824			
FT	DOMAIN 825 1021			
FT	DOMAIN 9 13			
FT	POLY-PRO.			
FT	POLY-ALA.			
FT	DOMAIN 14 17			

FT	DNA	20	24	POLY-GLU.
FT	DOMAIN	992	995	PHOSPHORYLATION (BY SIMILARITY).
FT	MOD. RES.	669	669	MISSING (IN REF. 3).
FT	CONFLICT	211	253	A -> P (IN REF. 2).
FT	CONFLICT	341	341	A -> P (IN REF. 3).
FT	CONFLICT	342	342	R -> T (IN REF. 1).
FT	CONFLICT	381	381	T -> S (IN REF. 1).
FT	CONFLICT	428	428	S -> T (IN REF. 1).
FT	CONFLICT	431	431	Q -> R (IN REF. 2 AND 3).
FT	CONFLICT	484	493	MISSING (IN REF. 2).
FT	CONFLICT	768	768	A -> R (IN REF. 1).
FT	CONFLICT	826	826	P -> A (IN REF. 3).
FT	CONFLICT	947	948	PT -> RA (IN REF. 3).
FT	CONFLICT	1023	1023	S -> A (IN REF. 2 AND 3).
FT	CONFLICT	1044	1044	V -> I (IN REF. 1).
SO	SEQUENCE	1135 AA;	127472 MW;	056554F9BACCAAD CRC64;
Query Match				
Best Local Similarity 88.2%; Score 30; DB 1; Length 1135;				
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0				
Oy	1	TGRIP 6	:	
Db	691	GRIIP 696		
RESULT 15				
ID	COAC_CHICK	STANDARD:	PRT:	2324 AA.
AC	P11029;			
DI	01-JUL-1989	(Rel. 11, Created)		
DT	01-JUL-1989	(Rel. 11, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin			
DE	carboxylase (EC 6.3.4.14)].			
GN	ACAC.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae;			
OX	NCHI_TaxID=9031;			
RA	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Liver;			
RA	MEDLINE=88133305; PubMed=2893793;			
RT	Takai T., Yokoyama C., Wada K., Tanabe T.;			
RT	"Primary structure of chicken liver acetyl-CoA carboxylase deduced			
RL	from cDNA sequence.";			
J	Biol. Chem. 263:2651-2657(1988).			
RP	SEQUENCE OF 493-820 FROM N.A.			
RC	TISSUE=Liver;			
RA	MEDLINE=87106011; PubMed=2879745;			
RT	Takai T., Wada K., Tanabe T.;			
RT	"Primary structure of the biotin-binding site of chicken liver			
RL	acetyl-CoA carboxylase.";			
F	Bess Lett. 212:98-102(1987).			
-1-	FUNCTION: CATALYZES THE RATE-LIMITING REACTION IN THE BIOGENESIS			
CC	OF LONG-CHAIN FATTY ACIDS; THIS PROTEIN CARRIES THREE FUNCTIONS:			
CC	BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND			
CC	CARBOXYLTRANSFERASE.			
-1-	CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate			
CC	+ malonyl-CoA.			
CC	-1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)			
CC	= ADP + phosphate + carboxybiotin-carboxyl-carrier protein.			
CC	-1- COFACTOR: BIOTIN.			
CC	-1- ENZYME REGULATION: BY PHOSPHORYLATION.			
CC	-1- PATHWAY: long-chain fatty acid biosynthesis; first (rate-limiting)			
CC	step			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-1- SIMILARITY: PARTIAL TO CARBAAMOYL PHOSPHATE SYNTHETASES.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: J03541; AAA48701.1; -
 DR EMBL: X05019; CAA28675.1; -
 DR PIR: A29924; A29924.
 DR HSSP: P24182; IDV1.
 DR InterPro: IPR001982; Biotin_attach.
 DR InterPro: IPR005482; Biotin_carb_C.
 DR InterPro: IPR000089; Biotin_lipoyl.
 DR InterPro: IPR000022; Carboxyl_trans.
 DR InterPro: IPR005479; Cpsase_L_D2.
 DR InterPro: IPR005481; Cpsase_L_N.
 DR Pfam: PF02785; Biotin_carb_C; 1.
 DR Pfam: PF00364; Biotin_lipoyl; 1.
 DR Pfam: PF01039; Carboxyl_trans; 1.
 DR Pfam: PF00289; Cpsase_L_chain; 1.
 DR Pfam: PF02786; Cpsase_L_D2; 1.
 DR PROSITE: PS00188; BIOTIN; 1.
 DR PROSITE: PS00866; CPSASE_1; 1.
 DR PROSITE: PS00867; CPSASE_2; 1.
 DR Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme;
 KW ATP-binding; Phosphorylation.
 FT NP_BIND 315 320 ATP (POTENTIAL).
 FT ACT_SITE 441 441 BY SIMILARITY.
 FT BINDING 786 786 BIOTIN.
 FT MOD_RES 78 78 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 80 80 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 1193 1193 PHOSPHORYLATION (BY SIMILARITY).
 FT DOMAIN 1936 1965 COENZYME A-BINDING (BY SIMILARITY).
 SQ SEQUENCE 2324 AA; 262717 MW; 3F1C341F01BBBF6 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 2324;
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
 :|||||
 Db 907 SGRIP 912

Search completed: August 20, 2003, 12:35:02
 Job time : 3.38822 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 11.3494 Seconds
(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-34

Perfect score: 34

Sequence: 1 TGRIPP 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	198	16	Q8EZU7
2	34	100.0	334	5	Q9VNU2
3	34	100.0	537	2	Q33417
4	34	100.0	999	10	Q82432
5	34	100.0	1008	10	Q9LKT6
6	34	100.0	1012	10	Q9LKT4
7	34	100.0	1016	10	Q8LHM9
8	34	100.0	1016	10	Q92KCS
9	33	97.1	364	11	Q8BXX3
10	33	97.1	395	11	Q8BXX3
11	33	97.1	423	11	Q8BXX3
12	33	97.1	442	11	Q8BXX3
13	33	97.1	518	10	Q8BL63
14	33	97.1	802	10	Q8L4U4
15	32	94.1	331	16	Q98L18
16	32	94.1	337	16	Q982C7

17	32	94.1	445	3	Q8X0W7
18	32	94.1	496	10	Q9SR26
19	32	94.1	505	10	Q94CC4
20	32	94.1	505	10	Q9M652
21	32	94.1	505	10	Q8H1P9
22	32	94.1	778	10	Q9AX30
23	31	91.2	81	5	Q8MS94
24	31	91.2	252	10	Q94L67
25	31	91.2	306	16	Q9UXY6
26	31	91.2	318	16	Q9JYV5
27	31	91.2	358	5	Q8MQ35
28	31	91.2	613	16	Q9ZJ51
29	31	91.2	615	16	Q9ZJ51
30	31	91.2	633	10	Q8LP72
31	31	91.2	671	5	P81538
32	31	91.2	717	5	Q9VW27
33	31	91.2	737	10	Q8S700
34	31	91.2	744	10	Q9SHU6
35	31	91.2	977	10	Q9C7T7
36	31	91.2	981	10	Q9M6A8
37	31	91.2	1002	10	Q9M6A8
38	31	91.2	1110	10	Q9E177
39	31	91.2	1141	10	Q9LHP4
40	31	91.2	1196	10	Q22476
41	31	91.2	1278	10	Q9AV65
42	31	91.2	1299	5	Q8MQ37
43	31	91.2	2427	5	Q8MQ36
44	30	88.2	85	5	Q27065
45	30	88.2	100	16	Q8CMA1

ALIGNMENTS

RESULT 1	Q8EZU7	PRELIMINARY;	PRT;	198 AA.
ID	Q8EZU7			
AC	Q8EZU7			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Signal peptidase I (EC 3.4.21.89).			
GN	SIPC OR IA3754.			
OS	Leptospira interrogans.			
OC	Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.			
OX	NCBI_TaxID=173;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;			
RA	Ren S.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AE011531; AN50952.1; -			
KW	Hydrolase; Complete proteome.			
SQ	SEQUENCE 198 AA; 22758 MW; 5DBB8F731B5237 CRC64;			
Query Match	100.0%; Score 34; DB 16; Length 198;			
Best Local Similarity	100.0%; Prod. No. 26;			
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 TGRIPP 6			
DB	158 TGRIPP 163			
RESULT 2	Q9VNU2	PRELIMINARY;	PRT;	334 AA.
ID	Q9VNU2			
AC	Q9VNU2			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	CG11440 protein (H101743p).			
GN	CG11440.			

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Abdayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mel M.-H., Ibeagwam C.,
 RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palczolo M., Pitman G.S., Pan S., Pollard J., Put V., Reese M.G.,
 RA Reihert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dreesen D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacle J., Pargass V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03597; AAF51828.1; -;
 DR EMBL: AY19567; AAM50221.1; -;
 DR FlyBase: FBgn0037163; CG11440.
 DR InterPro: IPR000326; PA_FTPase.
 DR Pfam: PF01569; PAP2; 1.
 SQ SEQUENCE 334 AA; 37845 MW; 5454C14C2908376 CRC64;

Query Match 100.0%; Score 34; DB 5; Length 334;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||||
 DB 231 TGRIPP 236

RESULT 3
 ID 033417 PRELIMINARY; PRT; 537 AA.
 AC 033417;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ABC-type transporter, putative membrane subunit.
 GN AtSR.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PAOI;
 RX MEDLINE=95262702; PubMed=7744061;
 RA Bell S., Kehrli H., James P., Staudemann W., Cook A.M., Leisinger T.,
 RA Kertesz M.A.;
 RT "Purification and characterization of the arylsulfatase synthesized by
 RT Pseudomonas aeruginosa PAO during growth in sulfate-free medium, and
 RT cloning of the arylsulfatase gene (atsA).";
 RL Eur. J. Biochem. 229:385-394(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PAOI;
 RX Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z48540; CAA88423.1; -;
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp; 1.
 SQ SEQUENCE 537 AA; 57823 MW; 5F8B669560AF81BB CRC64;

Query Match 100.0%; Score 34; DB 2; Length 537;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 |||||
 DB 279 TGRIPP 284

RESULT 4
 ID 082432 PRELIMINARY; PRT; 999 AA.
 AC 082432;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Leucine-rich receptor-like protein kinase.
 GN LRPK1.
 OS Malus domestica (Apple) (Malus sylvestris).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Rosales; Rosaceae; Maloideae; Malus.
 OX NCBI_TaxID=3750;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, Florida;
 RX MEDLINE=99454543; PubMed=10527419;
 RA Komjanc M., Festi S., Rizzotti L., Cattiveill L., Cervone F.,
 RA De Lorenzo G.;
 RT "A leucine-rich repeat receptor-like protein kinase (LRPK1) gene is
 RT induced in Malus x domestica by Venturia inaequalis infection and
 RT salicylic acid treatment.";
 RL Plant Mol. Biol. 40:945-957(1999).
 RN [2]
 RP SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF053127; AAC36318.1; -;
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_plant.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.

DR Pfam; PF00560; LRR; 18.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR PRINTS; PR00109; TYRKINASE.
 DR Pfam; PF00001; Prot_kinase; 1.
 DR PROSITE; PS05052; LRR_PS; 5.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
 KW Transferase
 SO SEQUENCE 999 AA; 110101 MW; 69965F9ED1E199F2 CRC64;

Query Match 100.0%; Score 34; DB 10; Length 999;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 DB 251 TGRIP 256

RESULT 5

O9LKZ6 PRELIMINARY; PRT; 1008 AA.
 AC O9LKZ6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Receptor-like protein kinase 1.
 GN RLK1.
 OS Glycine max (Soybean).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 RX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21363855; PubMed-11470843;
 RA Yamamoto E., Knapp H.T.;
 RT "Soybean receptor-like protein kinase genes: paralogous divergence of
 a gene family.";
 RL Mol. Biol. Evol. 18:1522-1531(2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF244888; AAF91322.1; -.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007090; LRR_plant.
 DR InterPro; IPR004838; NHRantf_1.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00560; LRR; 19.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00001; Prot_kinase; 1.
 DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
 DR PROSITE; PS05052; LRR_PS; 6.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 1008 AA; 109724 MW; 4C803E0F913C59E5 CRC64;

Query Match 100.0%; Score 34; DB 10; Length 1008;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 DB 484 TGRIP 489

RESULT 6

O9LKZ6

ID O9LKZ6 PRELIMINARY; PRT; 1012 AA.
 AC O9LKZ6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Receptor-like protein kinase 3.
 GN RLK3.
 OS Glycine max (Soybean).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 RX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21363855; PubMed-11470843;
 RA Yamamoto E., Knapp H.T.;
 RT "Soybean receptor-like protein kinase genes: paralogous divergence of
 a gene family.";
 RL Mol. Biol. Evol. 18:1522-1531(2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF244890; AAF91324.1; -.
 DR HSP; P12931; IFMK.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007090; LRR_plant.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR006025; Zn_MTpeptidase.
 DR Pfam; PF00560; LRR; 19.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRINTS; PR00109; TYRKINASE.
 DR PROSITE; PS05052; LRR_PS; 6.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SO SEQUENCE 1012 AA; 110322 MW; FDS55FE57F99815D CRC64;

Query Match 100.0%; Score 34; DB 10; Length 1012;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIP 6
 DB 489 TGRIP 494

RESULT 7

O8LHW9 PRELIMINARY; PRT; 1016 AA.
 AC O8LHW9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN P0022805.33.
 OS Oryza sativa (japonica cultivar-group).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzeae; Oryza.
 RX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
 clone:P0022805.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004262; BAC10827.1; -.
 DR Gramene; O8LHW9; -.
 DR InterPro; IPR001611; LRR.

```

DR InterPro: IPR007090; LRR_plant.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00560; LRR; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS50502; LRR_Ps; 5.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Transferase.
SQ SEQUENCE 1016 AA; 110741 MW; 533584DBC6EA66 CRC64;

Query Match
Best Local Similarity 100.0%; Score 34; DB 10; Length 1016;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 310 TGRVPP 315

RESULT 8
Q92KC5 PRELIMINARY; PRT; 83 AA.
AC Q92KC5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein R01632.
GN R01632 OR SMC00940.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396507; PubMed=1181430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Godiard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Goffeau T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle P., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591787; CAC46211.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 9007 MW; DFA51690ACF778C5 CRC64;

Query Match
Best Local Similarity 97.1%; Score 33; DB 16; Length 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 71 TGRVPP 76

RESULT 9
Q8BX33 PRELIMINARY; PRT; 364 AA.
AC Q8BX33;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE GPI transamidase homolog.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573(2002).
DR EMBL: AK043033; BAC31442.1;
SQ SEQUENCE 364 AA; 41039 MW; DED410BF2FAE7830 CRC64;

Query Match
Best Local Similarity 97.1%; Score 33; DB 11; Length 364;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 136 TGRVPP 141

RESULT 10
Q8BH63 PRELIMINARY; PRT; 395 AA.
AC Q8BH63;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE GPI transamidase homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573(2002).
DR EMBL: AK077893; BAC37051.1;
DR EMBL: AK082806; BAC38629.1;
SQ SEQUENCE 395 AA; 44895 MW; 4765ADDF4D46E386 CRC64;

Query Match
Best Local Similarity 97.1%; Score 33; DB 11; Length 395;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIPP 6
DB 136 TGRVPP 141

RESULT 11
Q8H7T3 PRELIMINARY; PRT; 423 AA.
AC Q8H7T3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative SKP1-like protein.
GN OSJNBA0009C08.20.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saeki C.,
 RA Currie J., Collura K.,
 RT "Rice Genomic Sequence."
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC107224; AAN60492.1;
 SO SEQUENCE 423 AA; 47756 MW; 2DE1454ACAC643E CRC64;

Query Match
 Best Local Similarity 97.1%; Score 33; DB 10; Length 423;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 10 TGRVPP 15

RESULT 12
 O8BL63 PRELIMINARY; PRT; 442 AA.

ID O8BL63;
 AC O8BL63;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Gpi transamidase homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK046246; BAC32653.1;
 SO SEQUENCE 442 AA; 49784 MW; 0A907F22B5230B0 CRC64;

Query Match
 Best Local Similarity 97.1%; Score 33; DB 11; Length 442;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 136 TGRVPP 141

RESULT 13
 O8R053 PRELIMINARY; PRT; 518 AA.

ID O8R053;
 AC O8R053;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Putative cytochrome P450.
 GN OSJNB0023M1.14.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrharioideae; Oryzaceae; Oryza.
 NC NCB1_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RA McCombe W.R., de la Bastide M., Spiegel L., Preston R., Kirchhoff K.,
 RA Kuit K., Nascento L., Baker J., Santos L., Zuttavern T., Miller B.,
 RA Cunniff D.M., Katzenberger F., Muller S., Bell M., Ballia V., Shah R.,
 RA King L., Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedhia N.,
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
 OSJNB0023M1.1, from chromosome 10, complete sequence."

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AC092749; AAM08560.1;
 DR Gramene: O8R053;
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SO SEQUENCE 518 AA; 57230 MW; 0D95E6BA9C4232 CRC64;

Query Match
 Best Local Similarity 97.1%; Score 33; DB 10; Length 518;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGRIPP 6
 DB 97 TGRVPP 102

RESULT 14
 O8L404 PRELIMINARY; PRT; 802 AA.

ID O8L404;
 AC O8L404;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative receptor-like protein kinase (Hypothetical protein).
 GN OSJNB003C23.12 OR OSJNB0038A07.18.
 OS Oryza sativa (Japanese cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrharioideae; Oryzaceae; Oryza.
 NC NCB1_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
 RA Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
 RA Uteback T.T., Feldblyum T.V., Yang Q.O., Haas B.J., Sun B.B.,
 RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 10 BAC OSJNB0038A07 genomic sequence."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
 RA Overton II L.L., Bera J.J., Teltin T., Krol M.I., Jarral B.B.,
 RA Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
 RA Uteback T.T., Feldblyum T.V., Yang Q.O., Haas B.J., Sun B.B.,
 RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 10 BAC OSJNB0038A07 genomic sequence."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC113948; AAM94518.1;
 DR Gramene: O8L404;
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR007090; LRR_Plant.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_Chr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR PRINTS: PRO0019; LEURICRPT.
 DR PRODOM: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYRKc; 1.
 DR PROSITE: PS05052; LRR_PS; 3.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; ATP-binding; Kinase; Receptor; Transferase.
 SO SEQUENCE 802 AA; 85525 MW; 41A650742C6C6144 CRC64;

Query Match
 Best Local Similarity 97.1%; Score 33; DB 10; Length 802;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
|||:|

Db 191 TGRVPP 196

RESULT 15

098LI8 PRELIMINARY; PRT; 331 AA.
ID 098LI8
AC 098LI8
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Transcriptional regulator.
GN MUR1005.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002996; BAB48475.1; -;
DR InterPro; IPR000843; HTH_LactI.
DR InterPro; IPR001761; PeriplABP/LactI.
DR Pfam; PF00532; Peripla_BP_1like; 1.
DR SMART; SM00354; HTH_LactI; 1.
KW Complete proteome.
SQ SEQUENCE 331 AA; 36699 MW; B50DDFF8798F9A2 CRC64;

Query Match

Best Local Similarity 94.1%; Score 32; DB 16; Length 331;
Pred. No. 1.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGRIP 6
|||:|

Db 311 TGRVPP 316

Search completed: August 20, 2003, 12:40:38
Job time : 14.4605 secs

DR WPI: 2000-039074/03.

XX Fibrinectin ED-B domain epitope specific antibodies and conjugate
PT antibodies

PS Claim 10; Page 38; 59pp; English.

XX The present sequence represents the VH component of a modified human
CC scFv antibody which has specific affinity for a characteristic epitope
CC of the ED-B domain of fibrinectin. The affinity of the antibody for
CC this epitope was improved by introducing a number of mutations in the
CC complementarity determining region (CDR) residues located at the
CC periphery of the binding site. The improved antibody is used for rapid
CC targeting markers of angiogenesis, for detecting diseases characterized
CC by vascular proliferation, such as diabetic retinopathy, age-related
CC macular degeneration or tumours. The antibody localizes the respective
CC tissue within 3 to 4 hours after injection. It is used in
CC immunoscientigraphic detection of angiogenesis and for diagnosis and
CC therapy of tumours and diseases characterized by vascular proliferation.
CC The antibody can be conjugated to a molecule which induces blood
CC coagulation and blood vessel occlusion. These conjugates are used in
CC the preparation of injectable compositions for the treatment of
CC angiogenesis-related pathologies, especially caused by or associated
CC with ocular angiogenesis.

XX Sequence 116 AA:

Query Match 100.0%; Score 608; DB 21; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.8e-49;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVRQAPGKGLRWSSISGSGTTY 60
DB 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVRQAPGKGLRWSSISGSGTTY 60
QY 61 ADSVKGRTISRNSKNTLYLQMSLRADTAVYYCAKFPFYDYGQGLTVTVSS 116
DB 61 ADSVKGRTISRNSKNTLYLQMSLRADTAVYYCAKFPFYDYGQGLTVTVSS 116

RESULT 2
AAE08817
ID AAE08817 standard; Protein; 116 AA.

XX AC AAE08817;

DT 19-NOV-2001 (first entry)

XX Human scFv L19 antibody variable heavy chain region.

XX Human; scFv; single-chain variable antibody fragment; cancer; cytotoxic;
KW coagulant; ED-B domain; fibrinectin; tumour; ocular disorder; psoriasis;
KW vascular proliferation; rheumatoid arthritis; blood vessel occlusion;
KW angiogenesis; blood coagulation; variable heavy chain; VH.

OS Homo sapiens.

PN WO200162800-A1.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001WO-EP02062.

PR 24-FEB-2000; 2000US-0512082.

PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.

PI Neri D, Tarli L, Viti F, Birchler M;

DR WPI: 2001-541701/60.

XX An antibody, with specific affinity for a characteristic epitope of the
PT ED-B domain of fibrinectin for the treatment of diseases characterized

PT by vascular proliferation

PS Claim 10; Page 36; 73pp; English.

XX The invention relates to an antibody with specific affinity for a
CC characteristic epitope of the ED-B domain of fibrinectin, where the
CC antibody has improved affinity to ED-B. The invention also relates to
CC conjugates comprising antibodies with a suitable photoreactive molecule
CC useful in the detection and/or coagulation of blood vessels. An antibody
CC with improved affinity to the ED-B domain is useful for diagnosis and
CC therapy of tumours and diseases characterised by vascular proliferation,
CC cancer, rheumatoid arthritis, neo-vasculature associated pathologies
CC disorders and psoriasis. Treatment of angiogenesis related pathologies
CC comprises the injection of conjugates comprising antibody and a molecule
CC capable of inducing blood coagulation and blood vessel occlusion. The
CC present sequence is single-chain variable antibody fragment (scFv) L19
XX antibody variable heavy chain (VH) region related to the invention.

XX Sequence 116 AA:

Query Match 100.0%; Score 608; DB 22; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.8e-49;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVRQAPGKGLRWSSISGSGTTY 60
DB 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFSFSMSWVRQAPGKGLRWSSISGSGTTY 60

QY 61 ADSVKGRTISRNSKNTLYLQMSLRADTAVYYCAKFPFYDYGQGLTVTVSS 116
DB 61 ADSVKGRTISRNSKNTLYLQMSLRADTAVYYCAKFPFYDYGQGLTVTVSS 116

RESULT 3
AAV53775
ID AAV53775 standard; Protein; 238 AA.

XX AC AAV53775;

DT 22-FEB-2000 (first entry)

XX An antibody with improved specificity for fibrinectin.

XX scFv; antibody; ED-B domain epitope; fibrinectin; marker;
KW angiogenesis; vascular proliferation; diabetic retinopathy;
KW age-related macular degeneration; tumour; immunoscientigraphic detection;
KW blood coagulation; blood vessel occlusion; ocular angiogenesis;
KW angiogenesis-related pathology.

OS Synthetic.

OS Homo sapiens.

PN WO9958570-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-EP03210.

PR 11-MAY-1998; 98US-0075338.

PR 28-APR-1999; 99US-0300425.

PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.

PI Neri D, Tarli L, Viti F, Birchler M;

DR WPI: 2000-039074/03.

XX Fibrinectin ED-B domain epitope specific antibodies and conjugate

PT antibodies
PS Claim 10; Page -; 59pp; English.
XX The present sequence represents a modified human scFv antibody which

XX WPI: 2001-451939/48.
DR N-PSDB: AAS22533.
XX
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
PS Example 4; Page 577; 894pp; English.

XX The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence represents a protein of the invention.

XX Sequence 363 AA;

Query Match 100.0%; Score 608; DB 22; Length 363;

Best Local Similarity 100.0%; Pred. No. 2e-48; Mismatches 0; Gaps 0;

Matches 116; Conservative 0; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVOPGGSRLSCAASGFTSSFSMSWVROAPGKLEWVSSISGSGTTY 60
DB 20 EVOLLESGGGLVOPGGSRLSCAASGFTSSFSMSWVROAPGKLEWVSSISGSGTTY 79

QY 61 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCAKPFYFDYWGOGTLVTVSS 116
DB 80 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCAKPFYFDYWGOGTLVTVSS 135

RESULT 6
ABR05053
ID ABR05053 standard; Protein; 117 AA.

XX ABR05053;

XX 26-MAR-2002 (first entry)

DE Fibronectin isoform B specific binding antibody VH domain protein.

XX Fibronectin isoform B; B-FN; extra domain B; Ed-B; antibody; VH domain;
KW angiogenesis; neoplasia; tumour; detection.

XX Unidentified.

XX WO200196599-A2.

XX 20-DEC-2001.

XX 08-JUN-2001; 2001WO-EP06533.

XX 15-JUN-2000; 2000US-211677P.

PA (PHIL-) PHILGEN SR.
PA (CAST/) CASTELLANI P.
PA (ZARD/) ZARDI L.
PA (ZIDL/) ZILJSTRA A.
XX Castellani P, Zardi L, Ziljstra A;

XX WPI: 2002-098076/13.
DR N-PSDB: ABA92717.

XX Detecting fibronectin isoform B in body fluid of an individual, by
PT binding fibronectin to gelatin on solid support, and determining
PT ability of specific binding member for Extra Domain B to be retained on
PT support -
PS Disclosure; Page 54; 72pp; English.

XX The present invention describes a method for detecting the presence of
CC fibronectin isoform B (B-FN) in body fluid of an individual. The method
CC involves passing the fluid over a solid support (SS) to which gelatin
CC that binds FN is attached, such that if FN is present, it is retained on
CC SS by binding gelatin, and determining the presence of B-FN on SS by
CC determining the ability of a specific binding member for extra domain B
CC (ED-B) of B-FN to be retained on SS. The method is useful for the
CC quantitative determination of B-FN in a body fluid taken or collected
CC from an individual. Other methods from the present invention can be used:
CC for quantitating the amount of an insoluble marker protein in a sample
CC of tissue or tumour; for determining the grade or severity of neoplasia
CC in a sample of tissue or tumour. An elevated level of B-FN in a body
CC fluid such as plasma, cerebral-spinal fluid or cystic fluid obtained
CC from an individual is indicative of the presence of a tumour in the
CC individual. The methods are useful in diagnostic and prognostic
CC situations. The methods can be readily applied to small biopsy samples
CC obtained in the clinic, thereby providing valuable information without
CC additional compromise to the patients or complications for the
CC clinicians. The present sequence represents the VH domain of a B-FN Ed-B
CC specific binding antibody molecule, which is used in the exemplification
CC of the present invention.

XX Sequence 117 AA;

Query Match 92.8%; Score 564.5; DB 23; Length 117;

Best Local Similarity 94.2%; Pred. No. 6.4e-45; Mismatches 0; Indels 7; Gaps 2;

Matches 113; Conservative 0; Mismatches 0; Indels 7; Gaps 2;

QY 1 EVOLLESGGGLVOPGGSRLSCAASGFTSSFSMSWVROAPGKLEWVSSISGSGTTY 60
DB 1 EVOLLESGGGLVOPGGSRLSCAASGFTSSFSMSWVROAPGKLEWVSSISGSGTTY 60

QY 61 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCAK---PFYFDYWGOGTLVTVSS 116
DB 61 ADSVKGRTISRDNKNTLYLQMNSLRAEDTAVYYCAKOKSAP---FDYWGOGTLVTVSS 117

RESULT 7
ABP95999
ID ABP95999 standard; Protein; 116 AA.

XX ABP95999;

XX 01-MAY-2003 (first entry)

DE HSA antibody related VH chain VH dummy.

XX Ligand: human serum albumin; HSA; antibody; cytostatic; anti-HIV;
KW antiinflammatory; antianaemic; immunosuppressive; neuroprotective;
KW dual-specific ligand; cancer; HIV infection; hepatitis; rubella;
KW anaemia; inflammation; autoimmune disorder; multiple sclerosis;
KW Crohn's disease; myasthenia gravis.

XX Homo sapiens.
OS Synthetic.

XX	PN	WO200302609-A2.
XX	PD	09-JAN-2003.
XX	PF	28-JUN-2002; 2002WO-GB03014.
XX	PR	28-JUN-2001; 2001GB-0015841.
XX	PA	(MEDI-) MEDICAL RES COUNCIL.
XX	PI	Winter G, Ignatovich O, Tomlinson I;
XX	DR	WPI: 2003-210246/20.
XX	PT	Dual-specific ligand having immunoglobulins with binding specificity to
PT	PT	different antigens or epitopes, useful for treating, preventing or
PT	PT	diagnosing diseases, e.g. cancer, HIV infection, inflammations, or
XX	PT	myasthenia gravis.
XX	XX	Example 1; Fig 3; 84pp; English.
XX	PS	The present invention describes a dual-specific ligand (I) comprising:
CC	CC	(a) a first single immunoglobulin variable domain with a binding
CC	CC	specificity to a first antigen or epitope; and (b) a second complementary
CC	CC	immunoglobulin single variable domain with a binding activity to a second
CC	CC	antigen or epitope. The binding domains are mutually complementary, and
CC	CC	the first and second domains lack mutually complementary domains that
CC	CC	share the same specificity. (I) has cytostatic, anti-HIV, antianaemic,
CC	CC	antiinflammatory, immunosuppressive and neuroprotective activities. The
CC	CC	dual-specific ligand is useful for treating, preventing or diagnosing
CC	CC	diseases, e.g. cancer, HIV infection, hepatitis, rubella, anaemia,
CC	CC	inflammations or autoimmune disorders (e.g. multiple sclerosis, Crohn's
CC	CC	disease or myasthenia gravis). The dual-specific ligand may be used to
CC	CC	reconstitute cytotoxic T-cells to a cancer cell. The dual-specific ligand is
CC	CC	also useful for monitoring the efficacy of drugs, as well as for
CC	CC	alleviating toxicity. The present sequence represents a human serum
CC	CC	albumin (HSA) related antibody VH sequence, which is used in an example
CC	CC	from the present invention.
XX	XX	Sequence 116 AA:
QY	Query Match	92.4%; Score 562; DB 24; Length 116;
DB	Best Local Similarity	92.2%; Pred. No. 1.1e-44;
Matches	107; Conservative	5; Mismatches 4; Indels 0; Gaps 0;
QY	1	EVQLLESGGGLVPGGSLRLSCAASGTFSSFSMSVROAPGKGLWYSSTISGSGCTTY 60
DB	1	EVQLLESGGGLVPGGSLRLSCAASGTFSSFSMSVROAPGKGLWYSSTISGSGCTTY 60
QY	61	ADSVNGRTTISDNSKNTFLYIQMNSLRAEDTAVYYCAKPFYFDYWGQGLTVYSS 116
DB	61	ADSVNGRTTISDNSKNTFLYIQMNSLRAEDTAVYYCAKSYGAFDYWGQGLTVYSS 116
XX	RESULT 8	
XX	AAV02472	
XX	AAV02472	standard; Protein: 240 AA.
XX	AAV02472;	
XX	15-JUL-1999	(first entry)
XX	DE	A single chain antibody (ScFv).
XX	KW	Screening: functional polypeptide; ligand; non-functional;
XX	KW	entitlement; single chain antibody; ScFv.
XX	OS	Unidentified.
XX	PN	WO9920749-A1.
XX	PD	29-APR-1999.

```

PE      20-OCT-1998;    98WO-GB03135.
PF      21-NOV-1997;    97US-0066729.
PR      20-OCT-1997;    97GB-0022131.
PR      13-NOV-1997;    97US-0065428.
XX
XX      (MEDI-) MEDICAL RES COUNCIL.
XX
XX      Tomlinson I, Winter G;
XX      WPI: 1999-288302/24.
DR      N-PSDB; AAX36070.
XX
XX      Screening for functional polypeptides which bind a ligand
PT
PS      Disclosure; Fig 2; 67pp; English.
XX
XX      The specification describes a method for screening for functional
CC      polypeptides which bind a ligand. The method comprises contacting a
CC      repertoire of polypeptides with a generic ligand, and then screening
CC      selected functional polypeptides with a target ligand. The method
CC      permits the removal from a chosen repertoire of polypeptides, those
CC      which are non-functional, e.g. as a result of the introduction of
CC      frame-shift mutations, stop codons, folding mutants or expression
CC      mutants, which would be or are incapable of binding to any target
CC      ligand. The method also permits the enrichment of a chosen repertoire
CC      of polypeptides for those polypeptides which are functional, well folded
CC      and highly expressed. The polypeptides obtained can be used in
CC      diagnostic, prophylactic and therapeutic procedures. The present
CC      sequence represents the single chain antibody (ScFv) that forms the
CC      basis of a library according to the invention.
SQ      Sequence    240 AA:
XX
XX      Query Match          92.4%; Score 562; DB 20; Length 240;
XX      Best Local Similarity 92.2%; Pred. No. 2.4e-44;
XX      Matches 107; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY      1 EVQLLESGGGLVPGGSLRLSCAASGFTFPSSMSWVRAPKGLGEWYSISGSSTYY 60
DQ      1 EVOQLLESGGGGLVPGGSLRLSCAASGFTFPSSYAMSVMWRAPKGLGEWSAISGSSTYY 60
DY      1 EVOQLLESGGGGLVPGGSLRLSCAASGFTFPSSYAMSVMWRAPKGLGEWSAISGSSTYY 60
DY      61 ADSVAGRFITSDNKNKNTLYLQMNLSLRADRTAVYYCAKPPFYFDYWGGTLVTYSS 116
DY      61 ADSVAGRFITSDNKNKNTLYLQMNLSLRADRTAVYYCAKSYGAIFYWGCGTLVTYSS 116
DY
DY      RESULT 9
DY      ABP95597
DY      ID ABP95597 standard; Protein; 240 AA.
DY
DY      ABP95597;
DY
DY      01-MAY-2003 (first entry)
DY
DY      Human serum albumin antibody related protein #1.
DE
DE      Ligand; human serum albumin; HSA; antibody; cytosstatic; anti-HIV;
DE      antinflammatory; antianaemic; immunosuppressive; neuroprotective;
DE      dual-specific ligand; cancer; HIV infection; hepatitis; rubella;
DE      anaemia; inflammation; autoimmune disorder; multiple sclerosis;
DE      Crohn's disease; myasthenia gravis.
OS      Homo sapiens.
OS      Synthetic.
XX
XX      WO2003002609-A2.
XX
XX      09-JAN-2003.
XX
XX      28-JUN-2002; 2002MO-GB03014.
XX
XX      28-JUN-2001; 2001GB-0015841.
XX

```

```

PA (MED-1) MEDICAL RES COUNCIL.
XX
PI Winter G, Ignatovich O, Tomlinson I;
XX WPI: 2003-210246/20.
DR N-PSDB; ABZ76706.
XX
XX Dual-specific ligand having immunoglobulins with binding specificity to
PT different antigens or epitopes, useful for treating, preventing or
PT diagnosing diseases, e.g. cancer, HIV infection, inflammations, or
PT myasthenia gravis
XX
XX Example 1; Fig 1; 84pp; English.
XX
XX The present invention describes a dual-specific ligand (I) comprising:
CC (a) a first single immunoglobulin variable domain with a binding
CC specificity to a first antigen or epitope; and (b) a second complementary
CC immunoglobulin single variable domain with a binding activity to a second
CC antigen or epitope. The binding domains are mutually complementary, and
CC the first and second domains lack mutually complementary domains that
CC share the same specificity. (I) has cytostatic, anti-HIV, antianaemic,
CC antiinflammatory, immunosuppressive and neuroprotective activities. The
CC dual-specific ligand is useful for treating, preventing or diagnosing
CC diseases, e.g. cancer, HIV infection, hepatitis, rubella, anaemia,
CC inflammations or autoimmune disorders (e.g. multiple sclerosis, Crohn's
CC disease or myasthenia gravis). The dual-specific ligand may be used to
CC recruit cytotoxic T-cells to a cancer cell. The dual-specific ligand is
CC also useful for monitoring the efficacy of drugs, as well as for
CC monitoring toxicity. The present sequence represents a human serum
CC albumin (HSA) related antibody sequence, which is used in an example
SO from the present invention.
XX
XX Sequence 240 AA:
XX
XX Query Match 92.4%; Score 562; DB 24; Length 240;
Best Local Similarity 92.2%; Pred. No. 2.4e-44;
Matches 107; Conservative 5; Mismatches 4; Indels 0; Gaps 0.
QY 1 EVQLLESGGGLVQPFGSLRLSCAASGFFTFSSSMKWQAQPEKGLEWYSTISGSSTYY 60
Dd 1 EVQLLESGGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPKGLIEWVAISGSGSYTY 60
QY 61 AHSVGRFTISRDNSKNTLYLQMNSLRADETAVYYCAKPFPPEDYWGGTFLTVASS 116
Dd 61 AHSVGRFTISRDNSKNTLYLQMNSLRADETAVYYCAKSYGAFDYWGOGTLTVASS 116
RESULTS
ID AUU14320
ID AUU14320 standard; Protein; 313 AA.
XX
XX AUU14320:
XX
XX 24-OCT-2001 (first entry)
XX
XX Human novel protein #191.
XX
XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
XX immunomodulatory; cytoskeletal; neuroprotective; vlnary; nocitopic;
XX anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
XX antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
XX thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
XX Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
XX tissue regeneration; immune disorder.
XX
XX Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-USO2623.

```

[illegible]

OS Homo sapiens.
 XX JP2002174635-A.
 XX 21-JUN-2002.
 XX 07-DEC-2000; 2000JP-0373259.
 XX 07-DEC-2000; 2000JP-0373259.
 XX (KAGU-) KAGAKU GIJYUNSU SHINKO JIGYODAN.
 XX (KOKU-) KOKURITSU SEISHIN SHINKETI CENT. SOCHO.
 XX WPI: 2002-579732/62.
 XX N-PSDB; ABK99213.
 XX Screening of an antibody used for the identification of an objective
 XX protein in high efficiency.
 XX Example 8; Fig 1; 43pp; Japanese.
 XX The invention relates to screening an antibody against a specific protein
 XX in an objective structure sample containing a protein group in high
 XX efficiency, comprising reacting an objective structure sample containing
 XX a protein group or a portion containing an objective protein in the
 XX sample with an antibody library, recovering the antibody combined with
 XX the protein, replicating the recovered antibody and reacting it with the
 XX objective protein at least once. The method is used for the
 XX identification of an objective protein. The present sequence is an
 XX antibody chain (derived from the DP47 heavy chain, the DP42 light
 XX chain) used in the method of the invention.
 XX Sequence 116 AA;
 SQ
 Query Match 91.8%; Score 558; DB 23; Length 116;
 Best Local Similarity 92.2%; Pred. No. 2.5e-44;
 Matches 107; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVROAPGKGLEWVSSISGSGSTYY 60
 DB 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVROAPGKGLEWVSSISGSGSTYY 60
 QY 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 DB 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 DB 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 RESULT 12
 AAO21548
 ID AAO21548 standard; Protein: 116 AA.
 AC AAO21548;
 DT 02-SEP-2002 (first entry)
 DE Antibody screening method related protein VH(DP-47).
 KW Screening antibody; 2-D electrophoresis; plural protein; protein spot;
 KW antibody library; proteomics.
 OS Unidentified.
 OS WO200242774-A1.
 PN 30-MAY-2002.
 PD 05-JUN-2001; 2001WO-JP04732.
 PF 24-NOV-2000; 2000JP-0358539.
 PR (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PA (NINA-) JAPAN NAT CENT NEUROLOGY & PSYCHIATRY.

PI Kaneko K;
 XX WPI: 2002-471742/50.
 DR N-PSDB; AAL39119.
 XX Screening an antibody using 2-D electrophoresis on plural proteins in
 XX samples for separating individual protein spots to react with an
 XX antibody library useful in proteomics and other biological sciences -
 XX Disclosure; Fig 1; 78pp; Japanese.
 XX The invention relates to a novel method for screening an antibody,
 XX comprising performing 2-D electrophoresis on plural proteins in a sample.
 XX Individual protein spots are separated by reacting them with an antibody
 XX library and then replicating the bound antibodies before reacting them
 XX with the spot proteins again. The method is useful for screening an
 XX antibody specific for a target protein, e.g. from a phage antibody
 XX library, which is useful in proteomics for studying various protein and
 XX complementary deoxyribonucleic acid (cDNA) expression libraries as well
 XX as gene functions, and in other biological and medical sciences. This
 XX sequence is a protein relating to the antibody screening method of the
 XX invention.
 XX Sequence 116 AA;
 SQ
 Query Match 91.8%; Score 558; DB 23; Length 116;
 Best Local Similarity 92.2%; Pred. No. 2.5e-44;
 Matches 107; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVROAPGKGLEWVSSISGSGSTYY 60
 DB 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVROAPGKGLEWVSSISGSGSTYY 60
 QY 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 DB 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 DB 61 ADSVKGRTISRNSKNTLYLQMSLRAEDTAVYYCAKPPPYDYGQGLTVYSS 116
 RESULT 13
 ABJ36939
 ID ABJ36939 standard; Protein: 177 AA.
 AC ABJ36939;
 DT 01-MAY-2003 (first entry)
 DE Anti-CD40 monoclonal antibody related protein SEQ ID No 64.
 KW Antiallergic; haemostatic; immunomodulator; cytosolic; antibody;
 KW human CD40; IL-12; LPS; lipopolysaccharide; IFNgamma; interferon gamma;
 KW dendritic cell; high G28-5; CD95 expression; high G28-5; B cell line;
 KW immunoadjuvant; anti-tumour agent; immunosuppressant; allergy;
 KW autoimmune disease; coagulation factor VIII inhibitor; anti-CD40.
 OS Unidentified.
 OS WO200288186-A1.
 PN 07-NOV-2002.
 PD 26-APR-2002; 2002WO-JP04292.
 PF 27-APR-2001; 2001WO-US13672.
 PR 11-MAY-2001; 2001JP-0142482.
 PR 05-OCT-2001; 2001JP-0310535.
 PR 26-OCT-2001; 2001US-0040244.
 XX (KIRI) KIRIN BEER KK.
 XX Miyakawa T, Yoshida H, Force WR, Chen X, Takahashi N;
 PI WPI, 2003-120463/11.
 DR N-PSDB; ABT31881.

XX Anti-CD40 monoclonal antibody with antagonist/agonist activity to CD40,
 PT or functional fragment, is useful in the treatment of e.g. autoimmune
 PT diseases or cancer
 XX
 PS Claim 15; Page 59; 94pp; Japanese.
 XX
 CC The invention relates to an antibody to human CD40, or its functional
 CC fragment, has at least one of the following properties: acting on
 CC dendritic cells to produce IL-12 in the presence of LPS
 CC (lipopolysaccharide) and IFNgamma (interferon gamma); acting on dendritic
 CC cells to activate maturation of the dendritic cells with high G28-5
 CC antibody; and activating CD95 expression with high G28-5 antibody against
 CC B cell line. Such antibodies or functional fragments can be used as
 CC immunoadjuvants, anti-tumour agents, immunosuppressants, and as remedies
 CC for autoimmune diseases, allergy or coagulation factor VIII inhibitors
 CC syndrome. This sequence represents a protein relating to the anti-CD40
 CC monoclonal antibody of the invention.
 CC
 SQ Sequence 177 AA;

Query Match 91.6%; Score 557; DB 24; Length 177;
 Best Local Similarity 87.1%; Pred. No. 4.9e-44;
 Matches 108; Conservative 5; Mismatches 3; Indels 8; Gaps 1;

OY 1 EVOLLSSGGGLVOPGSLRLSCAASGFTFSFMSWVRAPGKGLFWSSISGSGSTYY 60
 DB 20 EVOLLSSGGGLVOPGSLRLSCAASGFTFSFMSWVRAPGKGLFWSSISGSGSTYY 79
 OY 61 ADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAK-----PFYFDYWGQGLTV 112
 DB 80 ADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAKDGYYGSGSYGFYWGQGLTV 139
 OY 113 TVSS 116
 DB 140 TVSS 143

RESULT 14

ID AAB46042 standard; Peptide; 240 AA.

AC AAB46042;
 DT 23-MAR-2001 (first entry)

DE Human TF anti-idiotype antibody fragment H1.

XX MUC1; human; vaccine; conformation-dependent antigen; antibody; cancer;
 KM antiidiotypic antibody; cytostatic; virucidal; antibacterial; TF antigen;
 KM antiparasitic; infectious disease.
 XX

OS Homo sapiens.

PN WO200073430-A2.

PD 07-DEC-2000.

PF 29-MAY-2000; 2000WO-DE01809.

PR 27-MAY-1999; 99DE-1024405.

PR 09-SEP-1999; 99DE-1043016.

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

PI Goletz S, Karsten U;

DR WPI; 2001-049937/06.

XX Vaccines against conformation-dependent or non-peptide antigens, based
 PT on DNA encoding peptide which mimics the antigen, useful e.g. as
 PT antitumor vaccines
 XX

PS Disclosure; Page 11; 36pp; German.

XX This invention describes a novel vaccine (V1) against
 CC conformation-dependent antigens (CDA) comprising DNA (I) and/or an
 CC antibody, or peptide which immunologically imitates CDA, is new. (1)
 CC encodes a region of an antiidiotypic antibody (Ab2) or another peptide
 CC which: (a) specifically binds to the binding site of an antibody (Ab1)
 CC or an antigen binding molecule; and (b) immunologically mimics the
 CC initial antigen. The epitope is partially or completely
 CC conformation-dependent, and has an immunogenic structure defined by a
 CC specific spatial conformation of amino acids. (1) is used in the form
 CC of linear or circular naked DNA and/or with a viral vector and/or
 CC adjuvants. The products of the invention have cytostatic, virucidal,
 CC antibacterial and antiparasitic. The invention also describes (1) a
 CC corresponding vaccine (V2) against antigens which are not proteins or
 CC peptides, as defined above but which have epitopes which show an
 CC immunogenic structure; (2) preparing (V1) and (V2); (3) human
 CC antiidiotypic antibody fragments against the MUC1-conformation epitope
 CC having one of 31 approximately 60 residue amino acids sequences, all
 CC fully defined in the specification; (4) MUC1-conformation epitope mimics
 CC having one of 16 9-17 residue amino acid sequences, all fully in the
 CC specification; (5) antiidiotypic antibody fragments against the TF
 CC antigen having one of 24 approximately 200 residue amino acid sequences,
 CC fully defined in the specification; (6) TF carbohydrate epitope mimics
 CC having one of 25 7-13 residue amino acid sequences, all fully defined in
 CC the specification; and (7) DNA sequences encoding the fragments and
 CC derivatives defined in (3), 4, 5, or 6). (V1) and (V2) are used to treat
 CC cancer, and infectious diseases, e.g. caused by prions, viruses, bacteria
 CC and parasites. The vaccines are effective in cases where vaccination has
 CC previously not been possible.
 CC
 SQ Sequence 240 AA;

Query Match 90.6%; Score 551; DB 22; Length 240;
 Best Local Similarity 90.5%; Pred. No. 2.5e-43;
 Matches 105; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 EVOLLSSGGGLVOPGSLRLSCAASGFTFSFMSWVRAPGKGLFWSSISGSGSTYY 60
 DB 1 EVOLLSSGGGLVOPGSLRLSCAASGFTFSFMSWVRAPGKGLFWSSISGSGSTYY 60
 OY 61 ADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAKPFYFDYWGQGLTVYSS 116
 DB 61 ADSVKGRFTISRNSKNTLYLQMNSLRAEDTAVYCAKPFYFDYWGQGLTVYSS 116

RESULT 15

ID ABR01519 standard; Protein; 220 AA.

AC ABR01519;

DT 16-APR-2003 (first entry)

DE Human anti-TIMP-1 antibody heavy chain #17.

XX Human; antibody; tissue inhibitor of metalloproteinase-1; TIMP-1; VLCDR3;
 KM matrix metalloproteinase; MMP; variable heavy chain; VHCDR3; hepatocytic;
 KM variable light chain; cytosolic; nephrotropic; cardiac; liver fibrosis;
 KM alcoholic liver disease; cardiac fibrosis; acute coronary syndrome;
 KM lupus nephritis; glomerulosclerotic renal disease; lung cancer;
 KM idiopathic pulmonary fibrosis; benign prostate hypertrophy; colon cancer.
 XX

OS Homo sapiens.

PN WO200286085-A2.

PD 31-OCT-2002.

PF 24-APR-2002; 2002WO-US12801.

PR 24-APR-2001; 2001US-285683P.

PA (FARB) BAYER CORP.
PA (MORP-) MORPHOSYS AG.
XX

(MORP-) MORPHOSYS AG.

Pan C, Knorr AM, Schauer M, Hirth-dietrich C, Kraft S, Krebs B,

DR WPI; 2003-129114/12.
DR N-PSDB: AB274790

DR N-PSDB; AB274790
VY

PT new human anti-TIMP-1 (tissue inhibitor of metalloproteinase-1)
PT antibodies, for diagnosing or ameliorating the symptoms of a disorder
PT in which TIMP-1 is elevated, e.g., liver fibrosis, benign prostate
PT hypertrophy or lung cancer

PS Claim 20; page 159-160; 228pp; English
XX

The invention relates to a novel purified preparation of a human antibody, which binds to a tissue inhibitor of metalloproteinase-1 (TIMP-1), and neutralises a matrix metalloproteinase (MMP)-inhibiting activity of TIMP-1. The antibody comprises a variable heavy chain (VH)CD3 region and a variable light chain (VL)CD3 region. An antibody preparation of the invention has hepatotropic, cystostatic, nephrotropic and cardiant activity. The human antibody is useful for decreasing an MMP-inhibiting activity of a TIMP-1. It is especially useful for ameliorating the symptoms of a disorder in which TIMP-1 is elevated, e.g. liver fibrosis, alcoholic liver disease, cardiac fibrosis, acute coronary syndrome, lupus nephritis, glomerulosclerotic renal disease, idiopathic pulmonary fibrosis, benign prostate hypertrophy, lung cancer or colon cancer. The antibody is also useful for detecting a TIMP-1 in a test preparation, or in diagnosing a disorder in which a TIMP-1 level is elevated. The sequences shown in ABR01502-ABR01545 represent the heavy chain regions of a human anti-TIMP-1 antibody of the invention.

SQ Sequence 220 AA;

Query Match	Score	DB	Length
90.38;	549;	24;	220;
Best Local Similarity	89.79;	24;	220;

Matches	104;	Conservative	7;	Mismatches	5;	Indels	0;	Gaps	0
---------	------	--------------	----	------------	----	--------	----	------	---

QY

1 EVQLVDSGGGLVDPGGSRLRSCAASFTFSSFEFSMVNRAPGKGLEWVSISGSSTCY 60
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dd

1 QVQLVESGGGLVDPGGSRLRSCAASFTFSSYAMSMVNRAPGKGLEWVAISIGSGSTCY 60

1 VVVESSGGGVPGGSLKLSCAABGFTFSSYAMSWVRQAPGKGLEWVSAISGGSTYY 60

Dy 61 ADSYKGRFTISRDN SKNTLYLQNM SLRAEDTAVYYCAKPF EYFDYWGOGTLTVSS 116
| | | | | : | | | | |
Db 61 ADSYKGRFTISRDN SKNTLYLQNM SLRAEDTAVYYCARLG YFDLMGOGTLTVSS 116

01 ADSVKGREJ1SRDNSKNLTYLQMSLRADTAVYCCARLIGYFDLWGQGLTVSS 116

Search completed: August 20, 2003, 12:33:39
Job time : 256.06 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:40:51 ; Search time 125.783 Seconds

(without alignments)
121.698 Million cell updates/sec

Title: US-09-512-082-19

Perfect score: 608
Sequence: 1 EVQLLESGGGLVOPGSGRLTSCAASGFTFSFSMSWVAQAPGKGLIEWSSISGSGGTTY 116

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	608	100.0	116	11	US-09-300-425B-19
2	562	92.4	240	9	US-09-192-854-2
3	562	92.4	240	10	US-09-968-744A-2
4	562	92.4	240	12	US-09-968-744A-2
5	557	91.6	124	15	US-10-040-244-16
6	552.5	90.9	123	15	US-10-269-805-59
7	544	89.5	120	10	US-09-840-459-85
8	542.5	89.2	121	10	US-09-840-459-92
9	541.5	89.1	123	10	US-09-840-459-82
10	541	88.9	239	11	US-09-880-748-1922
11	540.5	88.9	123	15	US-10-269-805-23
12	540	88.8	288	10	US-09-818-247-22
13	539.5	88.7	125	10	US-09-840-459-76
14	539	88.7	120	14	US-10-025-687-4
15	539	88.7	120	15	US-10-125-687-4

16	537.5	88.4	113	9	US-09-056-160B-11	Sequence 11, Appl
17	537.5	88.4	224	11	US-09-972-656-68	Sequence 68, Appl
18	536	88.2	239	11	US-09-880-748-2018	Sequence 2018, Ap
19	534.5	87.9	449	10	US-09-736-371B-21	Sequence 21, Appl
20	534	87.8	116	10	US-09-840-459-80	Sequence 80, Appl
21	534	87.8	246	11	US-09-880-748-1975	Sequence 1975, Ap
22	534	87.8	247	11	US-09-880-748-1923	Sequence 1923, Ap
23	533	87.7	118	15	US-10-001-934-39	Sequence 39, Appl
24	533	87.7	239	11	US-09-880-748-2022	Sequence 2022, Ap
25	532.5	87.6	119	9	US-09-811-123-3	Sequence 3, Appl1
26	532.5	87.6	119	15	US-10-268-501-6	Sequence 6, Appl1
27	531.5	87.4	127	10	US-09-840-459-87	Sequence 87, Appl
28	531	87.3	239	11	US-09-880-748-2023	Sequence 2023, Ap
29	531	87.3	263	10	US-09-956-086-3	Sequence 3, Appl1
30	531	87.3	263	10	US-09-956-087-3	Sequence 3, Appl1
31	531	87.3	283	10	US-09-983-580-6	Sequence 6, Appl1
32	531	87.3	283	10	US-09-985-442-6	Sequence 6, Appl1
33	530	87.2	246	11	US-09-880-748-1980	Sequence 1980, Ap
34	529.5	87.1	130	15	US-10-118-100-54	Sequence 54, Appl
35	529.5	87.1	249	13	US-10-039-785-53	Sequence 53, Appl
36	528	86.8	128	10	US-09-840-459-77	Sequence 77, Appl
37	528	86.8	128	10	US-09-840-459-79	Sequence 79, Appl
38	528	86.8	250	11	US-09-840-459-84	Sequence 84, Appl
39	526.5	86.6	125	10	US-09-840-459-81	Sequence 81, Appl
40	525	86.3	124	10	US-09-880-748-2035	Sequence 2035, Ap
41	525	86.3	239	11	US-09-880-748-2035	Sequence 2035, Ap
42	525	86.3	443	9	US-09-917-410-4	Sequence 4, Appl1
43	524.5	86.3	130	15	US-10-118-100-53	Sequence 53, Appl
44	524.5	86.3	240	11	US-09-880-748-2013	Sequence 2013, Ap
45	524	86.2	249	11	US-09-880-748-1656	Sequence 1656, Ap

ALIGNMENTS

RESULT 1
US-09-300-425B-19
? Sequence 19, Application US/09300425B
? Publication NO. US20030045681A1
? GENERAL INFORMATION:
? APPLICANT: NERI, Dario
? APPLICANT: TARLI, Lorenzo
? APPLICANT: VITTI, Francesco
? APPLICANT: BIRCHLER, Manfred
? TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONTIGATES
? CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
? TITLE OF INVENTION: ANGIOGENESIS
? FILE REFERENCE: SCH-1733PI
? CURRENT APPLICATION NUMBER: US/09/300,425B
? CURRENT FILING DATE: 1999-04-28
? PRIOR APPLICATION NUMBER: 09/075,338
? PRIOR FILING DATE: 1998-05-11
? SOFTWARE: Patent Ver. 2.1
? SEQ ID NO 19
? LENGTH: 116
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: H antibody specific
? OTHER INFORMATION: for ED-B domain of fibronectin
US-09-300-425B-19

Query Match 100.0% Score 608; DB 11; Length 116;
Best Local Similarity 100.0%; Pred. No. 7.7e-47;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVQLLESGGGLVOPGSGRLTSCAASGFTFSFSMSWVAQAPGKGLIEWSSISGSGGTTY 60
DB 1 EVQLLESGGGLVOPGSGRLTSCAASGFTFSFSMSWVAQAPGKGLIEWSSISGSGGTTY 60
QY 61 ADVSGRTTISDKNKTLVQMSLRADPTAVYYCAKPFYFGGGLTVYSS 116
|||||

DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKPPFYEDYWGGLTVYSS 116

RESULT 2

US-09-192-854-2

Sequence 2, Application US/09192854

Patent No. US20020068276A1

GENERAL INFORMATION:

APPLICANT: Winter, Greg

APPLICANT: Tomlinson, Ian

TITLE OF INVENTION: Methods for Selecting Functional Peptides

FILE REFERENCE: 3789/72916

CURRENT APPLICATION NUMBER: US/09/192,854

CURRENT FILING DATE: 1998-11-17

EARLIER APPLICATION NUMBER: 60/066,729

EARLIER FILING DATE: 1997-11-21

NUMBER OF SEQ ID NOS: 212

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 2

LENGTH: 240

TYPE: PRT

ORGANISM: Homo sapiens

US-09-192-854-2

Query Match 92.4%; Score 562; DB 9; Length 240;

Best Local Similarity 92.2%; Pred. No. 1.9e-42;

Matches 107; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSFSSMSWVRQAPGKGLEWVSSISGSGTTY 60

DB 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSFSSYAMSVVRQAPGKGLEWVSAISGSGTTY 60

QY 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKPPFYEDYWGGLTVYSS 116

DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKSGAFDYWGGLTVYSS 116

RESULT 3

US-09-968-561A-2

Sequence 2, Application US/09968561A

Patent No. US2001064642A1

GENERAL INFORMATION:

APPLICANT: Tomlinson, Ian M

APPLICANT: Winter, Gregory

TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligands

FILE REFERENCE: 8039/1073B

CURRENT APPLICATION NUMBER: US/09/968,561A

CURRENT FILING DATE: 2001-10-01

PRIOR APPLICATION NUMBER: GB 9722131.1

PRIOR FILING DATE: 1997-10-20

PRIOR APPLICATION NUMBER: US 60/065,248

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: US 60/066,729

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: PCT/GB98/03135

PRIOR FILING DATE: 1998-10-20

PRIOR APPLICATION NUMBER: US 09/511,939

PRIOR FILING DATE: 2000-02-24

NUMBER OF SEQ ID NOS: 350

SOFTWARE: Patentln version 3.1

SEQ ID NO 2

LENGTH: 240

TYPE: PRT

ORGANISM: Homo sapiens

US-09-968-561A-2

Query Match 92.4%; Score 562; DB 10; Length 240;

Best Local Similarity 92.2%; Pred. No. 1.9e-42;

Matches 107; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSFSSMSWVRQAPGKGLEWVSSISGSGTTY 60

DB 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSYAMSVVRQAPGKGLEWVSAISGSGTTY 60

QY 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKPPFYEDYWGGLTVYSS 116

DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKSGAFDYWGGLTVYSS 116

RESULT 4

US-09-968-744A-2

Sequence 2, Application US/09968744A

Patent No. US20030148372A1

GENERAL INFORMATION:

APPLICANT: Tomlinson, Ian M

APPLICANT: Winter, Gregory

TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand

FILE REFERENCE: 8039/1073

CURRENT APPLICATION NUMBER: US/09/968,744A

CURRENT FILING DATE: 2003-01-13

PRIOR APPLICATION NUMBER: GB 9722131.1

PRIOR FILING DATE: 1997-10-20

PRIOR APPLICATION NUMBER: US 60/065,248

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: US 60/066,729

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: PCT/GB98/03135

PRIOR FILING DATE: 1998-10-20

PRIOR APPLICATION NUMBER: US 09/511,939

PRIOR FILING DATE: 2000-02-24

NUMBER OF SEQ ID NOS: 350

SOFTWARE: Patentln version 3.1

SEQ ID NO 2

LENGTH: 240

TYPE: PRT

ORGANISM: Homo sapiens

US-09-968-744A-2

Query Match 92.4%; Score 562; DB 12; Length 240;

Best Local Similarity 92.2%; Pred. No. 1.9e-42;

Matches 107; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSFSSMSWVRQAPGKGLEWVSSISGSGTTY 60

DB 1 EVOLLESGGGLVOPGGSLRLSCAASGFTFSYAMSVVRQAPGKGLEWVSAISGSGTTY 60

QY 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKPPFYEDYWGGLTVYSS 116

DB 61 ADVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKSGAFDYWGGLTVYSS 116

RESULT 5

US-10-040-244-16

Sequence 16, Application US/10040244

Patent No. US20030059427A1

GENERAL INFORMATION:

APPLICANT: KIRIN BEER KABUSHIKI KAISHA

APPLICANT: FORCE, WALKER F.

APPLICANT: TAKAHASHI, NOBUAKI

APPLICANT: MIKAYAMA, TOSHIFUMI

TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF HIGHLY ACTIVE ANTI-CD40 ANTI

FILE REFERENCE: 021286/0272501

CURRENT APPLICATION NUMBER: US/10/040,244

CURRENT FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: 60/200,601

PRIOR FILING DATE: 2000-4-28

PRIOR APPLICATION NUMBER: PCT/US01/13672

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: 09/844,684

PRIOR FILING DATE: 2001-04-27

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentln Ver. 3.0

SEQ ID NO 16

LENGTH: 124

TYPE: PRT

ORGANISM: Homo sapiens

US-10-040-244-16

Query Match 91.6%; Score 557; DB 15; Length 124;

Best Local Similarity 87.1%; Pred. No. 2.7e-42;

Matches 108; Conservative 5; Mismatches 3; Indels 8; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
DB 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
QY 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK-----PPYDYGQGLTV 112
|||||
DB 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKDGYYGSGGYGFDYWGQGLTV 120
|||||
QY 113 TVSS 116
|||||
DB 121 TVSS 124

RESULT 6

US-10-269-805-59

Sequence 59, Application US/10269805

Patent No. US20030124129A1

GENERAL INFORMATION:

APPLICANT: OLINER, JONATHAN D.

TITLE OF INVENTION: ANGIOPEPTIN-2 SPECIFIC BINDING AGENTS

FILE REFERENCE: A-722

CURRENT APPLICATION NUMBER: US/10/269,805

CURRENT FILING DATE: 2002-10-10

PRIOR APPLICATION NUMBER: US 60/328,604

PRIOR FILING DATE: 2001-10-11

NUMBER OF SEQ ID NOS: 76

SOFTWARE: PatentIn version 3.1

SEQ ID NO 59

LENGTH: 123

TYPE: PRT

ORGANISM: Homo sapiens

US-10-269-805-59

Query Match 90.9%; Score 552.5; DB 15; Length 123;

Best Local Similarity 87.8%; Pred. No. 6.7e-42;

Matches 108; Conservative 4; Mismatches 4; Indels 7; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
DB 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
QY 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK-----PPYDYGQGLTV 113
|||||
DB 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKETSFTFSYGFYDYGQGLTV 120
|||||
QY 114 VSS 116
|||||
DB 121 VSS 123

RESULT 7

US-09-840-459-85

Sequence 85, Application US/09840459

Patent No. US20020150576A1

GENERAL INFORMATION:

APPLICANT: Larosa, Gregory J.

APPLICANT: Horvath, Christopher

APPLICANT: Newman, Walter

APPLICANT: Jones, S. Tarran

APPLICANT: O'Brien, Stobhan H.

TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND

FILE REFERENCE: 1855.1052-012

CURRENT APPLICATION NUMBER: US/09/840,459

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: PCT/US01/03537

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/497,625

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: 09/359,193

PRIOR FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: 09/121,781

PRIOR FILING DATE: 1998-07-23

NUMBER OF SEQ ID NOS: 107

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 85

LENGTH: 120

TYPE: PRT

ORGANISM: Homo sapiens

US-09-840-459-85

Query Match 89.5%; Score 544; DB 10; Length 120;

Best Local Similarity 88.3%; Pred. No. 3.7e-41;

Matches 106; Conservative 5; Mismatches 5; Indels 4; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
DB 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
QY 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK-----PPYDYGQGLTV 116
|||||
DB 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKSPVYVPAADYWGQGLTV 120
|||||

RESULT 8

US-09-840-459-92

Sequence 92, Application US/09840459

Patent No. US20020150576A1

GENERAL INFORMATION:

APPLICANT: Larosa, Gregory J.

APPLICANT: Horvath, Christopher

APPLICANT: Newman, Walter

APPLICANT: Jones, S. Tarran

APPLICANT: O'Brien, Stobhan H.

TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND

FILE REFERENCE: 1855.1052-012

CURRENT APPLICATION NUMBER: US/09/840,459

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: PCT/US01/03537

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/497,625

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: 09/359,193

PRIOR FILING DATE: 1999-07-22

PRIOR APPLICATION NUMBER: 09/121,781

PRIOR FILING DATE: 1998-07-23

NUMBER OF SEQ ID NOS: 107

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 92

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

US-09-840-459-92

Query Match 89.2%; Score 542.5; DB 10; Length 121;

Best Local Similarity 86.8%; Pred. No. 5e-41;

Matches 105; Conservative 6; Mismatches 5; Indels 5; Gaps 1;

QY 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
DB 1 EVOLLESGGGLVOPGSLRLSCAAGFTFFSSFSMSWVQAAPGKLEWVSISGSGGTTY 60
|||||
QY 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKPPY-----FDYWGQGLTV 115
|||||
DB 61 ADVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKDISMDLEGLDYGQGLTV 120
|||||
QY 116 S 116

Db 121 \$ 121

RESULT 9

US-09-840-459-82
 ; Sequence 82, Application US/09840459
 ; Patent No. US20020150576A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Larosa, Gregory J.
 ; APPLICANT: Horvath, Christopher
 ; APPLICANT: Newman, Walter
 ; APPLICANT: Jones, S. Tarran
 ; APPLICANT: O'Brien, Siobhan H.
 ; APPLICANT: O'Keefe, Theresa
 ; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
 ; FILE REFERENCE: 1855.1052-012
 ; CURRENT APPLICATION NUMBER: US/09/840,459
 ; CURRENT FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: PCT/US01/03537
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 09/497,625
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR APPLICATION NUMBER: 09/359,193
 ; PRIOR FILING DATE: 1999-07-22
 ; PRIOR APPLICATION NUMBER: 09/121,781
 ; PRIOR FILING DATE: 1998-07-23
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 82
 ; LENGTH: 123
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-840-459-82

Query Match 89.1%; Score 541.5; DB 10; Length 123;
 Best Local Similarity 87.1%; Pred. No. 6,3e-41;
 Matches 108; Conservative 3; Mismatches 4; Indels 9; Gaps 2;

QY 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 Db 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 QY 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAKPP-----YFDYWGQGLTV 112
 Db 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAK-PGDYGSSTYLDYWGQGLTV 119
 QY 113 TVSS 116
 Db 120 TVSS 123

RESULT 10
 US-09-880-748-1922
 ; Sequence 1922, Application US/09880748
 ; Publication No. US20030059937A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunoselectively Bind Blys
 ; FILE REFERENCE: P5523
 ; CURRENT APPLICATION NUMBER: US/09/880,748
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25

NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1922
 ; LENGTH: 239
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-880-748-1922

Query Match 89.0%; Score 541; DB 11; Length 239;
 Best Local Similarity 88.8%; Pred. No. 1.4e-40;
 Matches 103; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 Db 1 EVOLVETGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 QY 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAKPPYFDYWGQGLTVSS 116
 Db 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAKRGVDYWGQGLTVSS 116

RESULT 11
 US-10-269-805-23
 ; Sequence 23, Application US/10269805
 ; Publication No. US20030124129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLINER, JONATHAN D.
 ; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
 ; FILE REFERENCE: A-722
 ; CURRENT APPLICATION NUMBER: US/10/269,805
 ; CURRENT FILING DATE: 2002-10-10
 ; PRIOR APPLICATION NUMBER: US 60/328,604
 ; PRIOR FILING DATE: 2001-10-11
 ; NUMBER OF SEQ ID NOS: 76
 ; SOFTWARE: PatentIn Version 3.1
 ; SEQ ID NO 23
 ; LENGTH: 123
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-269-805-23

Query Match 88.9%; Score 540.5; DB 15; Length 123;
 Best Local Similarity 85.4%; Pred. No. 7.7e-41;
 Matches 105; Conservative 6; Mismatches 5; Indels 7; Gaps 1;

QY 1 EVOLLESGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 Db 1 EVOLVDSGGGLVPGGSLRLSCAASGFTSSFSMSWVRQAPGKGLEWSSISGSGSTYY 60
 QY 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAKPPYFDYWGQGLTV 113
 Db 61 ADVKGRFTISRNSKNTLYLQMNSLRAEDTAVYYCAKETISFTSGYFDYWAQGLTV 120
 QY 114 VSS 116
 Db 121 VSS 123

RESULT 12
 US-09-818-247-22
 ; Sequence 22, Application US/09818247
 ; Patent No. US20020102657A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mostlov, Keith E.
 ; APPLICANT: Chaplin, Steven J.
 ; APPLICANT: Richman-Eisenstat, Janice
 ; TITLE OF INVENTION: The Regents of the University of California
 ; FILE REFERENCE: 18062E-000910US
 ; CURRENT APPLICATION NUMBER: US/09/818,247
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: WO PCT/US01/09669

```
;; PRIOR FILING DATE: 2001-03-26
;; PRIOR APPLICATION NUMBER: US 60/192,197
;; PRIOR FILING DATE: 2000-03-27
;; PRIOR APPLICATION NUMBER: US 60/192,198
;; PRIOR FILING DATE: 2000-03-27
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 22
;; LENGTH: 288
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial
US-09-818-247-22
```

```
Query Match      88.8%; Score 540; DB 10; Length 288;
Best Local Similarity 85.8%; Pred. No. 2e-40;
Matches 103; Conservative 9; Mismatches 4; Indels 4; Gaps 1;
```

```
QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
D 23 QVQLVSGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 82
QY 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFPF---YFDYWGQGLVTVSS 116
D 83 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKSFIVNSGYFQHMGGGLVTVSS 142
```

RESULT 13

```
US-09-840-459-76
;; Sequence 76, Application US/09840459
;; Patent No. US20020150576A1
;; GENERAL INFORMATION:
;; APPLICANT: Larosa, Gregory J.
;; APPLICANT: Horvath, Christopher
;; APPLICANT: Newman, Walter
;; APPLICANT: Jones, S. Tarran
;; APPLICANT: O'Brien, Siobhan H.
;; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
;; TITLE OF INVENTION: METHODS OF USE THEREFOR
;; FILE REFERENCE: 1855.1052-012
;; CURRENT APPLICATION NUMBER: US/09/840,459
;; PRIOR FILING DATE: 2001-02-02
;; PRIOR APPLICATION NUMBER: PCT/US01/03537
;; PRIOR FILING DATE: 2001-02-02
;; PRIOR APPLICATION NUMBER: 09/497,625
;; PRIOR FILING DATE: 2000-02-03
;; PRIOR APPLICATION NUMBER: 09/359,193
;; PRIOR FILING DATE: 1999-07-22
;; PRIOR APPLICATION NUMBER: 09/121,781
;; PRIOR FILING DATE: 1998-07-23
;; NUMBER OF SEQ ID NOS: 107
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 76
;; LENGTH: 125
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-840-459-76
```

```
Query Match      88.7%; Score 539.5; DB 10; Length 125;
Best Local Similarity 84.8%; Pred. No. 9.6e-41;
Matches 106; Conservative 6; Mismatches 4; Indels 9; Gaps 2;
```

```
QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
D 1 EVOLVESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
QY 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFPF---YFDYWGQGLVTVSS 111
D 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKIEPTAMFPYYIGMDVWGQGLTV 120
```

```
QY 112 VTWSS 116
D 121 VTWSS 125
```

RESULT 14

```
US-10-025-687-4
;; Sequence 4, Application US/10025687
;; Publication No. US20020142255A1
;; GENERAL INFORMATION:
;; APPLICANT: Luo, Peter
;; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
;; FILE REFERENCE: 26050-705
;; CURRENT APPLICATION NUMBER: US/10/025,687
;; CURRENT FILING DATE: 2002-04-17
;; NUMBER OF SEQ ID NOS: 28
;; SOFTWARE: Patent In version 3.1
;; SEQ ID NO 4
;; LENGTH: 120
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Human consensus antibody heavy chain variable region
US-10-025-687-4
```

```
Query Match      88.7%; Score 539; DB 14; Length 120;
Best Local Similarity 87.5%; Pred. No. 1e-40;
Matches 105; Conservative 6; Mismatches 5; Indels 4; Gaps 1;
```

```
QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
D 1 EVOLVESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
QY 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFPF---YFDYWGQGLVTVSS 116
D 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKRGDGFYAMDYWGQGLVTVSS 120
```

RESULT 15

```
US-10-125-687-4
;; Sequence 4, Application US/10125687
;; Publication No. US20030054407A1
;; GENERAL INFORMATION:
;; APPLICANT: Luo, Peter
;; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
;; FILE REFERENCE: 26050-705
;; CURRENT APPLICATION NUMBER: US/10/125,687
;; CURRENT FILING DATE: 2002-04-17
;; NUMBER OF SEQ ID NOS: 28
;; SOFTWARE: Patent In version 3.1
;; SEQ ID NO 4
;; LENGTH: 120
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Human consensus antibody heavy chain variable region
US-10-125-687-4
```

```
Query Match      88.7%; Score 539; DB 15; Length 120;
Best Local Similarity 87.5%; Pred. No. 1e-40;
Matches 105; Conservative 6; Mismatches 5; Indels 4; Gaps 1;
```

```
QY 1 EVOLLESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
D 1 EVOLVESGGGLVOPGGSLRLSCAAGFTFSFMSWVROAPGKLEWVSISGSGTYY 60
QY 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFPF---YFDYWGQGLVTVSS 116
D 61 ADVSKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKRWGDDGFYAMDYWGQGLVTVSS 120
```

Search completed: August 20, 2003, 13:16:44
Job time : 126.783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21; Search time 44.0241 seconds

(without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-19

Perfect score: 608
Sequence: 1 EVQLLESGGGLVPGGSLRL.....AKFPYFDYWGCGTLVTVSS 116

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483	79.4	117	1	P01764 mus sapien
2	470	77.3	136	1	P01765 mus musculu
3	469.5	77.2	115	1	P01766 mus musculu
4	466	76.6	122	1	P01767 mus sapien
5	462	76.0	114	1	P01768 mus sapien
6	447.5	73.6	115	1	P01769 mus sapien
7	444.5	73.1	120	1	P01770 mus sapien
8	444.5	73.1	121	1	P01771 mus sapien
9	444	73.0	122	1	P01772 mus sapien
10	443	72.9	116	1	P01773 mus sapien
11	434	71.4	122	1	P01774 mus sapien
12	433.5	71.3	116	1	P01775 mus sapien
13	433.5	71.3	119	1	P01776 mus sapien
14	430.5	70.8	117	1	P01777 mus sapien
15	429	70.6	119	1	P01778 mus sapien
16	429	70.6	119	1	P01779 mus sapien
17	429	70.6	119	1	P01780 mus sapien
18	429	70.6	119	1	P01781 mus sapien
19	429	70.6	119	1	P01782 mus sapien
20	426	70.1	114	1	P01783 mus sapien
21	426	70.1	126	1	P01784 mus sapien
22	424	69.7	119	1	P01785 mus sapien
23	423	69.6	117	1	P01786 mus sapien
24	418.5	68.8	115	1	P01787 mus sapien
25	418	68.8	111	1	P01788 mus sapien
26	418	68.8	122	1	P01789 mus sapien
27	417.5	68.7	118	1	P01790 mus sapien
28	416.5	68.5	113	1	P01791 mus sapien
29	416.5	68.4	122	1	P01792 mus sapien
30	414.5	68.2	119	1	P01793 mus sapien
31	414.5	68.2	123	1	P01794 mus sapien
32	414.5	68.2	123	1	P01795 mus sapien
33	414.5	68.2	123	1	P01796 mus sapien

34	411.5	67.7	113	1	P01797 mus sapien
35	411.5	67.7	123	1	P01798 mus sapien
36	410.5	67.5	117	1	P01799 mus sapien
37	410.5	67.5	117	1	P01800 mus sapien
38	410.5	67.5	123	1	P01801 mus sapien
39	410	67.4	119	1	P01802 mus sapien
40	408.5	67.2	113	1	P01803 mus sapien
41	408	67.1	117	1	P01804 mus sapien
42	407.5	67.0	115	1	P01805 mus sapien
43	407.5	67.0	116	1	P01806 mus sapien
44	407.5	67.0	116	1	P01807 mus sapien
45	407	66.9	117	1	P01808 mus sapien

ALIGNMENTS

RESULT 1

HY3C_HUMAN

ID HY3C_HUMAN STANDARD: PRT; 117 AA.

AC P01764;

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP MEDLINE=81101090; PubMed=6450418;

RA Matthyssens G., Rabbits T.H.;

RT "Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region."

RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).

CC -1-SIMILARITY: Contains 1 immunoglobulin-like domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC EMBL; J00236; AAA53516.1; -

CC EMBL; M35415; AAA58735.1; -

CC PIR; A02047; H3H026.

CC PDB; 1H0U; 23-DEC-99.

CC Genew; HGNC:5545; IGHE.

CC GO; GO:0005576; C:extracellular; NAS.

CC GO; GO:0003823; F:antigen binding activity; NAS.

CC GO; GO:0006955; F:immune response; NAS.

CC InterPro; IPR007110; Ig-like.

CC InterPro; IPR003006; Ig-MHC.

CC Pfam; PF00047; Ig_V.

CC SMART; SM00406; IGV_1.

CC PROSITE; PS00835; IG-Like; 1.

CC Immunoglobulin V region; Signal; 3D-structure.

CC STGNAL 1 19

CC CHAIN 20 117

CC DOMAIN 20 >117

CC NON_TER 117

CC SQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

CC IG HEAVY CHAIN V-III REGION VH26.

CC IG-LIKE.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 79.4%; Score 483; DB 1; Length 117;
Best Local Similarity 93.9%; Pred. No. 1.4e-41;
Matches 92; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

1 EVQLLESGGGLVPGGSLRLCAAGFTSSSMKVRQAPKGLKLEWSSISGGGCTTY 60

RA Lehman D.W., Putnam F.W.;
 RT "Amino acid sequence of the variable region of a human mu chain:
 RT location of a possible JH segment";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
 CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
 CC PATIENT WITH MACROGLOBULINEMIA.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR: A02051; M3HUM.
 DR HSSP: P01772; 2PB4.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS00835; IG-LIKE; 1.
 KW Immunoglobulin V region; Pyroliidone carboxylic acid.
 FT DOMAIN 1 112 IG-LIKE.
 FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
 FT NON_TER 122 122
 SQ SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

Query Match
 Best Local Similarity 76.6%; Score 466; DB 1; Length 122;
 Matches 88; Conservative 15; Mismatches 13; Indels 6; Gaps 1;

OY 1 EVOLLESGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 60
 DB 1 QVELVDSGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 60
 OY 61 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKPPYFDYWGQGLTVTV 114
 DB 61 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKPPYFDYWGQGLTVTV 120
 OY 115 SS 116
 DB 121 SS 122

RESULT 5
 HV3B_HUMAN STANDARD; PRT; 114 AA.
 ID P01763;
 AC 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region WEA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=83273707; PubMed=6410398;
 RA Goni F., Frangione B.;
 RT "Amino acid sequence of the Fv region of a human monoclonal IgM
 RT galactose in Klebsiella polysaccharides K30 and K33.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY
 CC AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH
 CC WALDENSTROM'S MACROGLOBULINEMIA.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR: A02046; M3HUM.
 DR HSSP: P01772; 2PB4.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS00835; IG-LIKE; 1.
 KW Immunoglobulin V region; Pyroliidone carboxylic acid.
 FT DOMAIN 1 112 IG-LIKE.
 FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
 FT NON_TER 114 114
 SQ SEQUENCE 114 AA; 12256 MW; D88294FBA18A07B7 CRC64;

Query Match
 Best Local Similarity 76.0%; Score 462; DB 1; Length 114;
 Matches 89; Conservative 14; Mismatches 11; Indels 2; Gaps 1;

OY 1 EVOLLESGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 60
 DB 1 QVELVDSGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 60
 OY 61 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKPPYFDYWGQGLTVTV 116
 DB 61 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAR--GWLNNQGLTVTVSS 114

RESULT 6
 HV3F_HUMAN STANDARD; PRT; 115 AA.
 ID P01767;
 AC 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-II region BUT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78137069; PubMed=416441;
 RA Torano A., Putnam F.W.;
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
 RT IgA2 immunoglobulin of the A2m (2) allotype.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
 CC -1- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C
 CC REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR: A02050; A2HUB.
 DR HSSP: P01769; 1MCP.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0003823; F:antigen binding activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; IGV_1.
 DR SMART: SM00406; IGV_1.
 DR PROSITE: PS00835; IG-LIKE; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 111 IG-LIKE.
 FT NON_TER 115 115
 SQ SEQUENCE 115 AA; 12379 MW; 208876A7DF52DCFA CRC64;

Query Match
 Best Local Similarity 73.6%; Score 447.5; DB 1; Length 115;
 Matches 87; Conservative 13; Mismatches 15; Indels 1; Gaps 1;

OY 1 EVOLLESGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 60
 DB 1 QVELVDSGGGLVOPGSLRLSCAASGFTFSFSMSWROAPGKLEWVSSISGSGSTYY 59
 OY 61 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKPPYFDYWGQGLTVTVSS 116
 DB 60 ADVSVGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKPPYFDYWGQGLTVTVSS 115

```

RESULT 7
HV3E_HUMAN          STANDARD;          PRT;          120 AA.
ID   HV3E_HUMAN
AC   P01766;
DT   21-JUL-1986 (Rel. 01, Created)
DR   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region BRO.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=7117674; PubMed=65324;
RT   Capra J.D., Hopper J.E.;
RT   "Comparative studies on monocytic Igm lambda and Igg kappa from an
RT   individual patient. III. The complete amino acid sequence of the VH
RT   region of the Igm paraprotein."
RL   Immunochimistry 13:995-999(1976).
CC   -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE
CC   SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM
CC   TYPE.
CC   -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02049; M3HUB4.
DR   HSSP; P01772; 2PB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro: IPR007110; Ig-like.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; IGV; 1.
DR   PROSITE; PS50835; IG-LIKE; 1.
KW   Immunoglobulin V region.
FT   DOMAIN 1 111 IG-LIKE.
FT   NON_TER 120 120
SQ   SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

Query Match          73.1%; Score 444.5; DB 1; Length 120;
Best Local Similarity 72.7%; Pred. NO. 1e-37;
Matches 88; Conservative 11; Mismatches 11; Indels 11; Gaps 2;

QY 1 EVOLLESGGLVOPGSLRLSCAASGFTFSFSMSWVROAPGKLEWVSSISGSGTTY 60
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1 EVOLVESGGGLVOPGSLRLSCAASGFTFSYNNWVROVYTKGLEWVSAI-GTAGDOY 59
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 ADVKGRFTISDNKNTLYIQMNSLRADPAVYVCAKPF-----PYFDYWGQGLT 110
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 60 ADVKGRFTISDNKNTLYIQMNSLRADPAVYVCAKSPVSLVDGMLYIYIGSVGQGT 119
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 111 L 111
DB 120 L 120

RESULT 8
HV3J_HUMAN          STANDARD;          PRT;          121 AA.
ID   HV3J_HUMAN
AC   P01771;
DT   21-JUL-1986 (Rel. 01, Created)
DR   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region HIL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=79124695; PubMed=420800;
RX   Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;

```

```

RT "Amino acid sequence of the VH region of human myeloma
RT cryoimmunoglobulin Igg H11."
RL Biochemistry 18:553-560(1979).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02054; G1H0HL.
DR   HSSP; P01772; 2PB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro: IPR007110; Ig-like.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; IGV; 1.
DR   PROSITE; PS50835; IG-LIKE; 1.
KW   Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT   DOMAIN 1 112 IG-LIKE.
FT   MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT   NON_TER 121 121
SQ   SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;

Query Match          73.1%; Score 444.5; DB 1; Length 121;
Best Local Similarity 70.5%; Pred. NO. 1e-37;
Matches 86; Conservative 13; Mismatches 16; Indels 7; Gaps 2;

QY 1 EVOLLESGGLVOPGSLRLSCAASGFTFSFSMSWVROAPGKLEWVSSISGSGTTY 60
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1 QVKLVQAGGAVQVQPSRLSLRCIASGFTFSYNGMHWVROAPGKLEWVAVIMYGSRTY 60
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 ADVKGRFTISDNKNTLYIQMNSLRADPAVYVCAKPF-----YFDYWGQGLT 114
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 61 GDSVGRFTISDNKNTLYIQMNSLRADPAVYVCAKPF-----YFDYWGQGLT 119
    :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 115 SS 116
DB 120 SS 121

RESULT 9
HV3A_HUMAN          STANDARD;          PRT;          122 AA.
ID   HV3A_HUMAN
AC   P01762;
DT   21-JUL-1986 (Rel. 01, Created)
DR   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ig heavy chain V-III region TRO.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE (MYELOMA PROTEIN TRO).
RX   MEDLINE=76023781; PubMed=809331;
RX   Kratzin H., Altevogt P., Ruban E., Kortt A., Starosciak K.,
RX   Hilschmann N.;
RT   "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.),
RT   II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT   structure of the complete IgA-molecule."
RL   Hoppe-Seyler's Z. Physiol. Chem 356:1337-1342(1975).
CC   -1- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.
CC   -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR   PIR; A02045; A1HUTR.
DR   HSSP; P01772; 2PB4.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0003823; F:antigen binding activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro: IPR007110; Ig-like.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; IGV; 1.

```

DR	PROSITE: PS50835; IG_LIKE; 1.
KW	Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT	DOMAIN 1 108 IG_LIKE.
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID
FT	NON_TER 122 122
SO	SEQUENCE 122 AA; 13472 MW; 2E21A11DA04DB0F9 CRC64;

Query Match	73.08;	Score 444;	DB 1;	Length 122;
Best Local Similarity	65.68;	Pred. No. 1.2e-37;		
Matches	80;	Conservative	18;	Mismatches 18;
			Indels	6;
			Gaps	1;

[illegible]

	RESULT	10
1	HVJ3T_HUMAN	
2	AC	
3	ID	
4	HVJ3T_HUMAN	
5	P01781:	STANDARD; PRT; 116 AA.
6	21-JUL-1986 (Rel. 01, Created)	
7	21-JUL-1986 (Rel. 01, Last sequence update)	
8	15-SEP-2003 (Rel. 42, Last annotation update)	
9	Ig heavy chain V-II region GAL.	
10	Hom sapiens (Human).	
11	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
12	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
13	NCBI_TaxId=9606;	
14	[1]	
15	RP	
16	SEQUENCE.	
17	MEDLINE=75059123; PubMed=4803843;	
18	Watanabe S., Banrikol H.U., Horn J., Bertram J., Hirschmann N.;	
19	"The primary structure of a monoclonal IgM-immunoglobulin	
20	(macroglobulin Gal.) II: the amino acid sequence of the H-chain (mu	
21	type), subgroup H III. Architecture of the complete IgM-molecule."	
22	Hope-Seyler's Z. Physiol. Chem.354:1505-1509(1973).	
23	[2]	

RP REVISION TO 28-33.
RA Hilschmann N.;
CC Submitted (JUN-1975) to the PIR data bank.
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC PIR: A02064; M3HNC.

DR	HSSP; P01772; F8B4.
DR	GO; GO:0005576; C:extracellular; NAS.
DR	GO; GO:0003983; F:antigen binding activity; NAS.
DR	GO; GO:0006925; P:immune response; NAS.
DR	InterPro: IPR007110; Ig-like.
DR	InterPro: IPR003006; Ig_MHC.
DR	InterPro: IPR003596; Ig_v.
DR	Pfam: PF00047; Ig; 1.
DR	SMART; SM00406; IGV; 1.
DR	PROSITE; PS50835; IG_LIKE; 1.
DR	Immunoglobulin V region.
CDR1	DOMAIN
CDR2	NON_TER
CDR3	1
CDR4	116
CDR5	116
CDR6	IG-LIKE.
SEQUENCE	116 AA; 12730 MW; 2C67CA9AAAAA1282 CRC64;

Query Match	72.9%;	Score 443;	DB 1;	Length 116;
Best Local Similarity	73.3%;	Pred. No. 1.4e-37;		
Matches	85;	Conservative	10;	Mismatches 21; Indels

1 EVQLLESGGGLVQPGGSLRLSCASGFTFSFSMSWVRQAPGKGLEWYSSISGSGTTY 60
|||:|||| |||| ||||||||| | :|||||||||||:::| |

Db 1 EVQLVESGGGLVQPGSRSLRLSCAASGFBFBELMTWRQAPGKGLELVANIKRBSGZ28Y 60

QY 61 ADSVKGRFTISRDNSSKNTLYLQNSLRLAEEDTAVYYCAKPPYFDYWGCGTLTVSS 116

61 VDSVKGFTISRDNKNSLYLQNSLRVEDTALYYCARGGGGDDYWGCGTLTVST 116

RESULT 11	
HV3H_HUMAN	
ID	STANDARD;
HV3H_HUMAN	PRT; 122 AA.
NC	
P01769.	

DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, last sequence update)
15-SEP-2003 (rel. 42, last annotation update)
19 heavy chain V-III region GA.
DE Homo sapiens (Human).
OS
DS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
11

RX	MEDLINE=741755307; Pubmed=4208643;
RA	Florent G., Jehman D., Putnam F.W.;
RT	"The switch point in mu heavy chains of human IgM immunoglobulins,"
RL	Biochemistry 13:2482-2498(1974).
CC	-I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S MACROGLOBULIN.
CC	-I- SIMILARITY: Contains 1 immunoglobulin-like domain.
PDR	PIR: A02052; M3HUGA.
DR	HSP: P01772; ZPB4.
DR	GO: GO:0005576; C:extracellular; NAS.
DR	GO: GO:0003823; F:antigen binding activity; NAS.
DR	GO: GO:0006935; F:immune response; NAS.
DR	InterPro: IPR007110; Ig-like.
DR	InterPro: IPR003006; IG_MMC.
DR	InterPro: IPR003596; Ig_v.
DR	pfam: PF00047; Ig; 1.
DR	SMART: SMO0406; IGV; 1.
DR	PROSITE: PS50835; IG_LIKE; 1.
KK	Immunoglobulin V region; Pyrroldione carboxylic acid.
FT	DOMAIN 1 112
FT	MOD_RES 1 1
FT	NON_TER 122 122
SEQUENCE	122 AA; 13166 MW; 74ESB6959EB4100A CRC64; IG-LIKE. PYRROLIDONE CARBOXYLIC ACID.

Query Match	71.4%;	Score 434;	DB 1;	Length 122;
Best Local Similarity	65.6%;	Pred. No. 1.2e-36;		
Matches	80;	Conservative	20;	Mismatches 16;
				Indels 6;
				Gaps 1

[illegible]

RESULT 12	
HV05_CARAU	
ID HV05_CARAU	STANDARD;
P10181.	PRT; 116 AA

01-NOV-1990 (Rel. 16, Created)
01-NOV-1990 (Rel. 16, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
15 heavy chain V region 5A precursor.
Carassius auratus (Goldfish).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

CC Cyrtinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8814476; PubMed=3125551;
 RA Wilson M.R., Middleton D., Warr G.W.;
 RT "Immunoglobulin heavy chain variable region gene evolution: structure
 RT and family relationships of two genes and a pseudogene in a teleost
 RT fish.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
 DR HSSP; P01772; 2FB4.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig; 1.
 DR SMART; SM00406; IgV; 1.
 DR PROSITE; PS50835; Ig_LIKE; 1.
 DR Immunoglobulin V region; Signal.
 FT CHAIN 1 116 IG HEAVY CHAIN V REGION 5A.
 FT SIGNAL 19
 FT DOMAIN 20 49 FRAMEWORK-1.
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 55 68 FRAMEWORK-2.
 FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 85 116 FRAMEWORK-3.
 FT DISULFID 41 114 BY SIMILARITY.
 FT NON_TER 116
 SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;
 Query Match 71.3%; Score 433.5; DB 1; Length 116;
 Best Local Similarity 87.8%; Pred. No. 1.2e-36;
 Matches 86; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
 QY 1 EVQLLESGGGLVQPQGSRLRSCASGTFPSFSMSVWRQAPGKLEWVSSISGSGSTYY 60
 DB 20 EVQLVDSGGGLVQPGSRLRSCASGTFYSSNYMWRQPPKGLFWVSVIR-SGGSITY 78
 QY 61 ADSVKGRTISRDNKNTLYLQMSLRADPTAVYCYAR 98
 DB 79 ADSVKGRTISRDNKNTLYLQMSLRADPTAVYCYAR 116
 RESULT 13
 H31_HUMAN STANDARD; PRT; 119 AA.
 ID HV31_HUMAN
 AC P01770;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-III region NIE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE.
 RX MEDLINE=77070269; PubMed=826475;
 RA Ponting H., Hilschmann N.;
 RT "The rule of antibody structure. The primary structure of a
 RT monoclonal IgG1 immunoglobulin (myeloma protein Nle). III. The
 RT chymotryptic peptides of the H-chain, alignment of the tryptic
 RT peptides and discussion of the complete structure.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
 RP DISULFIDE BOND.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nle). I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT characterizations of the products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA

CC PROTEIN.
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC PIR; A91668; G1HUN1.
 DR HSSP; P01772; 2FB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IgV; 1.
 DR PROSITE; PS50835; Ig_LIKE; 1.
 DR Immunoglobulin V region; Pyroglutamate carboxylic acid.
 FT CHAIN 1 112 IG-LIKE.
 FT MOD_RES 1 112 PYROGLUTAMATE CARBOXYLIC ACID.
 FT DISULFID 22 96
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 13242 MW; C96935A6E55E165B CRC64;
 Query Match 71.3%; Score 433.5; DB 1; Length 119;
 Best Local Similarity 70.6%; Pred. No. 1.3e-36;
 Matches 84; Conservative 15; Mismatches 17; Indels 3; Gaps 1;
 QY 1 EVQLLESGGGLVQPQGSRLRSCASGTFPSFSMSVWRQAPGKLEWVSSISGSGSTYY 60
 DB 1 EVQLVDSGGGLVQPGSRLRSCASGTFYSSNYMWRQPPKGLFWVSVIR-SGGSITY 78
 QY 61 ADSVKGRTISRDNKNTLYLQMSLRADPTAVYCYAR---PPYFDYMGQGLTVYSS 116
 DB 61 ADSVKGRTISRDNKNTLYLQMSLRADPTAVYCYARINDTAMFAMHQGLTVYSS 119
 RESULT 14
 H32_CANFA STANDARD; PRT; 117 AA.
 ID HV02_CANFA
 AC P01785;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V region MOO.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 OX [1]
 RP SEQUENCE OF 1-112.
 RX MEDLINE=77242268; PubMed=407924;
 RA Wasserman R.L., Capra J.D.;
 RT "Primary structure of the variable regions of two canine
 RT immunoglobulin heavy chains.";
 RL Biochemistry 16:3160-3168(1977).
 RN [2]
 RP SEQUENCE OF 113-117.
 RX MEDLINE=80077682; PubMed=117299;
 RA McCumber L.J., Capra J.D.;
 RT "The complete amino-acid sequence of a canine mu chain.";
 RL Mol. Immunol. 16:565-570(1979).
 CC -I- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
 CC -I- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC PIR; A90403; MHD60.
 DR HSSP; P01772; 2FB4.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IgV; 1.
 DR PROSITE; PS50835; Ig_LIKE; 1.
 DR Immunoglobulin V region.
 FT CHAIN 1 116 IG-LIKE.
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 12703 MW; FCE07309E0A84B35 CRC64;

Query Match 70.8%; Score 430.5; DB 1; Length 117;
 Best Local Similarity 73.3%; Pred. No. 2.5e-36;
 Matches 88; Conservative 11; Mismatches 14; Indels 7; Gaps 3;

OY 1 EVOLLESGGGLVOPGSGRLSCAASGFTFSFSMSWVROAPGKGLEWVSSISGSSGTTY 60
 DB 1 EVKLVESSGGDLVKGPGSGRLSCVASGFTFSNGMSWVRODPGEGLOWVADIS--SSGTTY 59
 OY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK----PPPYFDYWGQGLTVVSS 116
 DB 60 ADAVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCATGEGDIEIP--RYFGQGLTVVSS 117

RESULT 15

HV37_MOUSE
 ID HV37_MOUSE STANDARD; PRT; 119 AA.
 AC P01807;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 1g heavy chain V region X44.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]

SEQUENCE
 RX MEDLINE=79223895; PubMed=111245;
 RA Rao D.N., Rudnikoff S., Kruttsch H., Potter M.;
 RT "Structural evidence for independent joining region gene in
 RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
 RT its potential role in generating diversity in
 RT complementarity-determining regions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894(1979).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGA MYELOMA PROTEIN
 CC THAT BINDS GALACTAN.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC PIR: A02077; AYMSX4.
 DR HSSP: P01810; 2FBJ.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig: 1.
 DR SMART: SM00406; IGV: 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 DR Immunoglobulin V region.
 KW DOMAIN
 FT NON_TER 1 117 IG-LIKE.
 FT SEQUENCE 119 AA; 13246 MW; BC34FC8F31CD41B3 CRC64;

Query Match 70.6%; Score 429; DB 1; Length 119;
 Best Local Similarity 69.5%; Pred. No. 3.6e-36;
 Matches 82; Conservative 15; Mismatches 19; Indels 2; Gaps 1;

OY 1 EVOLLESGGGLVOPGSGRLSCAASGFTFSFSMSWVROAPGKGLEWVSSISGSSGTTY 60
 DB 1 EVKLVESSGGGLVOPGSGRLSCAASGFTFSMSWVROAPGKGLEWIGETINPDSSSTINY 60
 OY 61 ADVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK--PPPYFDYWGQGLTVVSS 116
 DB 61 TPSLNDKFTISRDNKNTLYLQMSKVRSEDTALYYCARLHYGYAAVWGQGLTVVSA 118

Search completed: August 20, 2003, 12:34:48
 Job time : 44.1352 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 219,422 seconds

(without alignments)
136,423 Million cell updates/sec

Title: US-09-512-082-19

Perfect score: 608
Sequence: 1 EVQLLESGGGLVQPGGSLRL.....AKPFYFDYWGQGLVTVSS 116

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	525.5	86.4	597	4	096BB9
2	510	83.9	471	4	08TC77
3	500.5	82.3	121	4	09UL71
4	494	81.2	118	4	09UL91
5	491.5	80.8	113	4	09UL90
6	484	79.6	613	4	08WUK1
7	475	78.1	112	4	08WUC1
8	473	77.8	118	4	08UJ72
9	472	77.6	147	4	09Y509
10	469	77.1	373	4	08WU38
11	468	77.0	132	4	09UL84
12	466.5	76.7	116	4	09UL93
13	463	76.2	494	4	096K68
14	458.5	75.4	119	11	0920E7
15	456.5	75.1	499	4	08NSK4
16	456	75.0	487	11	099KA4

17	452	74.3	473	11	091Z05	091Z05 mus musculu
18	443.5	72.9	479	11	091WP5	091WP5 mus musculu
19	438.5	72.1	131	4	09UL88	09UL88 homo sapien
20	432.5	71.1	493	4	08NCL6	08NCL6 homo sapien
21	427.5	70.3	486	11	091Z07	091Z07 mus musculu
22	420	69.1	469	11	08R3V9	08R3V9 mus musculu
23	418	68.8	95	4	09ULB6	09ULB6 homo sapien
24	417.5	68.7	480	11	091XE1	091XE1 mus musculu
25	411	67.6	437	11	09R1A4	09R1A4 mus musculu
26	403	66.3	484	11	08VER0	08VER0 mus musculu
27	400	65.8	298	11	09QXF0	09QXF0 mus musculu
28	398.5	65.5	112	4	09UGF3	09UGF3 homo sapien
29	398.5	65.5	521	4	08N4Y9	08N4Y9 homo sapien
30	392.5	64.6	104	4	09UL87	09UL87 homo sapien
31	386	63.5	124	4	09UL92	09UL92 homo sapien
32	370	60.9	124	6	09N0M6	09N0M6 homo sapien
33	369	60.7	124	6	09N0M4	09N0M4 oryctolagus
34	353.5	58.1	473	11	09DBL4	09DBL4 mus musculu
35	346.5	57.0	125	4	09UL95	09UL95 homo sapien
36	345	56.7	463	11	09QXC4	09QXC4 mus musculu
37	343.5	56.5	117	11	09QXF0	09QXF0 mus musculu
38	343	56.4	241	11	0921A6	0921A6 mus musculu
39	342.5	56.3	119	4	09UL94	09UL94 mus musculu
40	340.5	56.0	146	11	0924R8	0924R8 homo sapien
41	340	55.9	116	4	09UL89	09UL89 mus musculu
42	339	55.8	613	11	08VCX7	08VCX7 mus musculu
43	337.5	55.5	117	11	09QXK9	09QXK9 mus musculu
44	337	55.4	145	11	0924Q7	0924Q7 mus musculu
45	336.5	55.3	484	11	099LA6	099LA6 mus musculu

ALIGNMENTS

RESULT 1

ID Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC015760; ANH15760.1; -
DR InterPro: IPR007110; I9-11ke.
DR InterPro: IPR003006; I9-MHC.
DR InterPro: IPR003596; I9-V.
DR Pfam: PF00047; I9-5.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00835; IG-LIKE; 5.
DR PROSITE: PS00290; IG-MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FC3AD8BCE263D9 CRC64;

Query Match 86.4%; Score 525.5; DB 4; Length 597;
Best Local Similarity 83.2%; Pred. No. 8.2e-46;
Matches 104; Conservative 8; Mismatches 4; Indels 9; Gaps 2;

QY	1	EVQLLESGGGLVQPGGSLRLCAASGFTFSFSMSWVROAPGKGLIEWSSISGSGGTY	60
DB	20	EVQLLESGGGLVQPGGSLRLCAASGFTFSFSYAMWVROAPGKGLIEWSSISGSGGTY	79
QY	61	ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAK-PPPYF-----DYWGQGL	111
DB	80	ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAKPPGYSASGNTYREDYWGQGL	139


```

RESULT 5
O9UL90          PRELIMINARY:      PRT:      113 AA.
AC O9UL90:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
   (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
   "Myosin-reactive autoantibodies in rheumatic carditis and normal
   fetus.";
RT Clin. Immunol. Immunopathol. 87:184-192(1998).
RL EMBL: AF035024; AAD56260.1; -.
DR HSSP: P01772; 2PB4.
DR InterPro: IPR007110; Ig-1Ike.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PSS0835; IG-LIKE; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match      80.8%; Score 491.5; DB 4; Length 113;
Best Local Similarity 82.8%; Pred. No. 3.2e-43;
Matches 96; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 1 EVOLLESGGGLVOPGSGSLRLSCAASGFTFSFSMSWVRQAPGKGLEWVSSISGSGGTTY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVOLVESGGGVVOPGSGSLRLSCAASGFTFSYGMHWVRQAPGKGLEWVAIFIRYGSNKYY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFPDYWGQGLTVVSS 116
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFPDYWGQGLTVVSS 113
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
O8WUK1          PRELIMINARY:      PRT:      613 AA.
AC O8WUK1:
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=tonsil;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC020240; AAH20240.1; -.
DR InterPro: IPR007110; Ig-1Ike.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PSS0835; IG-LIKE; 5.
DR NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match      79.6%; Score 484; DB 4; Length 613;
Best Local Similarity 79.2%; Pred. No. 1.7e-41;
Matches 95; Conservative 7; Mismatches 14; Indels 4; Gaps 1;

QY 1 EVOLLESGGGLVOPGSGSLRLSCAASGFTFSFSMSWVRQAPGKGLEWVSSISGSGGTTY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 QVQLVESGGGVVOPGSGSLRLSCAASGFTFSYGMHWVRQAPGKGLEWVAISTYGSNKYY 79
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPF---PYDPYWGQGLTVVSS 116
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 80 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKADSEGVETPDWQGMVTVSS 139
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
O9HCC1          PRELIMINARY:      PRT:      112 AA.
AC O9HCC1:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Kikuchi M., Takada C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme: Isolation from a
   human synthetic phage display library and characterization.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB049915; BAB16829.1; -.
DR HSSP: P01772; 2PB4.
DR InterPro: IPR007110; Ig-1Ike.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
DR PROSITE: PSS0835; IG-LIKE; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match      78.1%; Score 475; DB 4; Length 112;
Best Local Similarity 79.5%; Pred. No. 1.6e-41;
Matches 89; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 EVOLLESGGGLVOPGSGSLRLSCAASGFTFSFSMSWVRQAPGKGLEWVSSISGSGGTTY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVOLVESGGGVVOPGSGSLRLSCAASGFTFDYGMHWVRQAPGKGLEWVSGIMNMGSGTGY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKPPFPDYWGQGLTV 112
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADVKGRFTISRDNKNTLYLQNNSLRAEDTAVYYCAKRRYALDYWGQGLTV 112
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
O9UL72          PRELIMINARY:      PRT:      118 AA.
AC O9UL72:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
   (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```


DR EMBL: AF035030; AAD56266.1; -
 DR HSSP: P01772; 2F84.
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 FT NON_TER 1
 FT 122 122
 SQ SEQUENCE 122 AA; 13579 MW; 36054DA1366545B8 CRC64;

Query Match 77.0%; Score 468; DB 4; Length 122;
 Best Local Similarity 75.4%; Pred. No. 9.5e-41;
 Matches 92; Conservative 8; Mismatches 16; Indels 6; Gaps 1;

OY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSVNRAPKGLGWSSISGSGTYY 60
 DB 1 EVOLVSGGGVOPGSLRLSCAASGFTSSNGMHVNRAPKGLGWAAISNDGSKNFY 60
 OY 61 ADSVKGRTISRDNKNTLYLQMSLRADTAIVYCAK-----FPYDWGGTLYTV 114
 DB 61 ADSVKGRTIFRDNKSNMMDLQMSLRADTAIVYCADEGRGLVGTIFDWGGTLYTV 120
 OY 115 SS 116
 DB 121 SS 122

RESULT 12
 O9UL93 PRELIMINARY; PRT: 116 AA.
 AC O9UL93;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2003 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Myosin-reactive immunoglobulin heavy chain variable region (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035021; AAD56257.1; -
 DR HSSP: P01772; 2F84.
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 1.
 FT NON_TER 1
 FT 116 116
 SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154D6061 CRC64;

Query Match 76.7%; Score 466.5; DB 4; Length 116;
 Best Local Similarity 80.2%; Pred. No. 1.3e-40;
 Matches 93; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

OY 2 VOLLESGGGLVOPGSLRLSCAASGFTSSFSMSVNRAPKGLGWSSISGSGTYYA 61
 DB 1 VOLVESGGGVOPGSLRLSCAASGFTSSYAMHVNRAPKGLGLEWVAISIDGSKNYA 60
 OY 62 DSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKPPY-FDYWGCGTLYTVSS 116
 DB 61 DSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAAGGGGLGIDYWGCGTLYTVSS 116

RESULT 13

O96K68 PRELIMINARY; PRT: 494 AA.

AC O96K68;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ14473.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
 RA Wagatsuma M., Hosolai T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kiwura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagehara K., Masuno T.,
 RA Nimomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AK027379; BAB5072.1; -
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 4
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PSS0835; IG_LIKE; 4.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 494 AA; 53088 MW; 9A1D7AB5AE4C0E CRC64;

Query Match 76.2%; Score 463; DB 4; Length 494;
 Best Local Similarity 75.4%; Pred. No. 1.9e-39;
 Matches 92; Conservative 11; Mismatches 13; Indels 6; Gaps 2;

OY 1 EVOLLESGGGLVOPGSLRLSCAASGFTSSFSMSVNRAPKGLGWSSISGSGTYY 60
 DB 20 EVOLVSGGGLVOPGSLRLSCAASGLSTYAMHVNRAPKGLGWSSISRDYIY 79
 OY 61 ADSVKGRTISRDNKNTLYLQMSLRADTAIVYCAK-----PPY-FDYWGCGTLYTV 114
 DB 80 RDSVKGRTISRDNKNTLYLQMSLRVDTAIVYCAKRDSCNGAICYGFSPMGCGTLYTV 139
 OY 115 SS 116
 DB 140 SS 141

RESULT 14

O920E7 PRELIMINARY; PRT: 119 AA.

AC O920E7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Pterin-mimicking anti-Idiotope heavy chain variable region (fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Atkin J.D., Iape A., Jennings I.G., Horafis O., Cotton R.G.H.;
 RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed
 in Mammalian Cells";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

Search completed: August 20, 2003, 12:40:11
Job time : 221.533 secs

DR EMBL: AF307937; AA109421.1; -
DR InterPro: IPR007110; IG_Like.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
FT NON_TER 1 1
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;

Query Match 75.4%; Score 456.5; DB 11; Length 119;
Best Local Similarity 73.9%; Pred. No. 8.8e-40;
Matches 88; Conservative 13; Mismatches 15; Indels 3; Gaps 1;

QY 1 EVQLLESGGGLVOPGSGLRSCAAGFTFSFSMSWVRQAPGKLEWVSSISGSGTYY 60
DB 1 EVQLLESGGGLVOPGSGLRSCAAGFTFSFSMSWVRQAPGKLEWVSSISGSGTYY 60
QY 61 ADSVKGRTISRDNKNTLYLQMSLRADTAAYYCAKPPY---FDYWGQTLVTVSS 116
DB 61 PDSVKGRTISRDNKNTLYLQMSLRADTAAYYCAKPPY---FDYWGQTLVTVSS 116

RESULT 15

Q8N5K4 PRELIMINARY; PRT; 499 AA.
ID 2 Q8N5K4
AC Q8N5K4;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC032249; AAH32249.1; -
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG_Like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; IG; 4.
DR SMART: SM00409; IG; 4.
DR SMART: SM00407; IGV; 2.
DR PROSITE: PS50835; IG_LIKE; 4.
DR PROSITE: PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 499 AA; 53376 MW; 93A5C89582054F32 CRC64;

Query Match 75.1%; Score 456.5; DB 4; Length 499;
Best Local Similarity 70.3%; Pred. No. 8.9e-39;
Matches 90; Conservative 11; Mismatches 14; Indels 13; Gaps 2;

QY 1 EVQLLESGGGLVOPGSGLRSCAAGFTFSFSMSWVRQAPGKLEWVSSISGSGTYY 60
DB 20 EVQLLESGGGLVOPGSGLRSCAAGFTFSFSMSWVRQAPGKLEWVSSISGSGTYY 79
QY 61 ADSVKGRTISRDNKNTLYLQMSLRADTAAYYCAKPPY---FDYWGQ 108
DB 80 ADSVKGRTISRDNKNTLYLQMSLRADTAAYYCAKPPY---FDYWGQ 108
QY 109 GTLVTVSS 116
DB 139 GTLVTVSS 146

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:35:06 ; Search time 143.289 Seconds

(without alignments)
85.038 Million cell updates/sec

Title: US-09-512-082-20

Perfect score: 73

Sequence: 1 GDCSSGSGGASTG 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents, AA, Main: *

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	14	14	US-09-075-338C-20
2	73	100.0	14	17	US-09-300-425B-20

3	73	100.0	14	19	US-09-512-082-20	Sequence 20, Appl
4	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78345</td><td>Sequence 78345, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78345</td> <td>Sequence 78345, A</td>	238	22	US-09-791-537-78345	Sequence 78345, A
5	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78361</td><td>Sequence 78361, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78361</td> <td>Sequence 78361, A</td>	238	22	US-09-791-537-78361	Sequence 78361, A
6	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78362</td><td>Sequence 78362, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78362</td> <td>Sequence 78362, A</td>	238	22	US-09-791-537-78362	Sequence 78362, A
7	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78363</td><td>Sequence 78363, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78363</td> <td>Sequence 78363, A</td>	238	22	US-09-791-537-78363	Sequence 78363, A
8	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78364</td><td>Sequence 78364, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78364</td> <td>Sequence 78364, A</td>	238	22	US-09-791-537-78364	Sequence 78364, A
9	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78365</td><td>Sequence 78365, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78365</td> <td>Sequence 78365, A</td>	238	22	US-09-791-537-78365	Sequence 78365, A
10	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78366</td><td>Sequence 78366, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78366</td> <td>Sequence 78366, A</td>	238	22	US-09-791-537-78366	Sequence 78366, A
11	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78367</td><td>Sequence 78367, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78367</td> <td>Sequence 78367, A</td>	238	22	US-09-791-537-78367	Sequence 78367, A
12	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78368</td><td>Sequence 78368, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78368</td> <td>Sequence 78368, A</td>	238	22	US-09-791-537-78368	Sequence 78368, A
13	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78369</td><td>Sequence 78369, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78369</td> <td>Sequence 78369, A</td>	238	22	US-09-791-537-78369	Sequence 78369, A
14	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78370</td><td>Sequence 78370, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78370</td> <td>Sequence 78370, A</td>	238	22	US-09-791-537-78370	Sequence 78370, A
15	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78371</td><td>Sequence 78371, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78371</td> <td>Sequence 78371, A</td>	238	22	US-09-791-537-78371	Sequence 78371, A
16	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78372</td><td>Sequence 78372, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78372</td> <td>Sequence 78372, A</td>	238	22	US-09-791-537-78372	Sequence 78372, A
17	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78373</td><td>Sequence 78373, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78373</td> <td>Sequence 78373, A</td>	238	22	US-09-791-537-78373	Sequence 78373, A
18	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78374</td><td>Sequence 78374, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78374</td> <td>Sequence 78374, A</td>	238	22	US-09-791-537-78374	Sequence 78374, A
19	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78375</td><td>Sequence 78375, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78375</td> <td>Sequence 78375, A</td>	238	22	US-09-791-537-78375	Sequence 78375, A
20	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78376</td><td>Sequence 78376, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78376</td> <td>Sequence 78376, A</td>	238	22	US-09-791-537-78376	Sequence 78376, A
21	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78377</td><td>Sequence 78377, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78377</td> <td>Sequence 78377, A</td>	238	22	US-09-791-537-78377	Sequence 78377, A
22	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78378</td><td>Sequence 78378, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78378</td> <td>Sequence 78378, A</td>	238	22	US-09-791-537-78378	Sequence 78378, A
23	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78379</td><td>Sequence 78379, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78379</td> <td>Sequence 78379, A</td>	238	22	US-09-791-537-78379	Sequence 78379, A
24	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78380</td><td>Sequence 78380, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78380</td> <td>Sequence 78380, A</td>	238	22	US-09-791-537-78380	Sequence 78380, A
25	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78381</td><td>Sequence 78381, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78381</td> <td>Sequence 78381, A</td>	238	22	US-09-791-537-78381	Sequence 78381, A
26	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78382</td><td>Sequence 78382, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78382</td> <td>Sequence 78382, A</td>	238	22	US-09-791-537-78382	Sequence 78382, A
27	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78383</td><td>Sequence 78383, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78383</td> <td>Sequence 78383, A</td>	238	22	US-09-791-537-78383	Sequence 78383, A
28	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78384</td><td>Sequence 78384, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78384</td> <td>Sequence 78384, A</td>	238	22	US-09-791-537-78384	Sequence 78384, A
29	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78385</td><td>Sequence 78385, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78385</td> <td>Sequence 78385, A</td>	238	22	US-09-791-537-78385	Sequence 78385, A
30	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78386</td><td>Sequence 78386, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78386</td> <td>Sequence 78386, A</td>	238	22	US-09-791-537-78386	Sequence 78386, A
31	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78387</td><td>Sequence 78387, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78387</td> <td>Sequence 78387, A</td>	238	22	US-09-791-537-78387	Sequence 78387, A
32	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78388</td><td>Sequence 78388, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78388</td> <td>Sequence 78388, A</td>	238	22	US-09-791-537-78388	Sequence 78388, A
33	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78389</td><td>Sequence 78389, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78389</td> <td>Sequence 78389, A</td>	238	22	US-09-791-537-78389	Sequence 78389, A
34	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78390</td><td>Sequence 78390, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78390</td> <td>Sequence 78390, A</td>	238	22	US-09-791-537-78390	Sequence 78390, A
35	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78391</td><td>Sequence 78391, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78391</td> <td>Sequence 78391, A</td>	238	22	US-09-791-537-78391	Sequence 78391, A
36	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78392</td><td>Sequence 78392, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78392</td> <td>Sequence 78392, A</td>	238	22	US-09-791-537-78392	Sequence 78392, A
37	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78393</td><td>Sequence 78393, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78393</td> <td>Sequence 78393, A</td>	238	22	US-09-791-537-78393	Sequence 78393, A
38	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78394</td><td>Sequence 78394, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78394</td> <td>Sequence 78394, A</td>	238	22	US-09-791-537-78394	Sequence 78394, A
39	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78395</td><td>Sequence 78395, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78395</td> <td>Sequence 78395, A</td>	238	22	US-09-791-537-78395	Sequence 78395, A
40	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78396</td><td>Sequence 78396, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78396</td> <td>Sequence 78396, A</td>	238	22	US-09-791-537-78396	Sequence 78396, A
41	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78397</td><td>Sequence 78397, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78397</td> <td>Sequence 78397, A</td>	238	22	US-09-791-537-78397	Sequence 78397, A
42	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78398</td><td>Sequence 78398, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78398</td> <td>Sequence 78398, A</td>	238	22	US-09-791-537-78398	Sequence 78398, A
43	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78399</td><td>Sequence 78399, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78399</td> <td>Sequence 78399, A</td>	238	22	US-09-791-537-78399	Sequence 78399, A
44	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78400</td><td>Sequence 78400, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78400</td> <td>Sequence 78400, A</td>	238	22	US-09-791-537-78400	Sequence 78400, A
45	73 <td>100.0<td>238</td><td>22</td><td>US-09-791-537-78401</td><td>Sequence 78401, A</td></td>	100.0 <td>238</td> <td>22</td> <td>US-09-791-537-78401</td> <td>Sequence 78401, A</td>	238	22	US-09-791-537-78401	Sequence 78401, A

ALIGNMENTS

RESULT 1
US-09-075-338C-20
Sequence 20, Application US/09075338C
GENERAL INFORMATION:
APPLICANT: NERI, Dario
APPLICANT: TARLI, Lorenzo
APPLICANT: VITTI, Francesco
APPLICANT: BIRCHIERI, Manfred
TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY
FILE REFERENCE: SCH-1733
CURRENT APPLICATION NUMBER: US/09/075, 338C
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: antibody linker
US-09-075-338C-20

Query Match 100.0%; Score 73; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. NO. 0.04; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

QY 1 GDSSGGSGGASTG 14
|||||
DB 1 GDSSGGSGGASTG 14

RESULT 2

US-09-300-425B-20
; Sequence 20, Application US/09300425B
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P1
; CURRENT APPLICATION NUMBER: US/09/300,425B
; CURRENT FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 20
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antibody linker
US-09-300-425B-20

Query Match 100.0%; Score 73; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.04; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0

QY 1 GDSSGGSGGASTG 14
|||||
DB 1 GDSSGGSGGASTG 14

RESULT 3

US-09-512-082-20
; Sequence 20, Application US/09512082
; GENERAL INFORMATION:
; APPLICANT: NERI, Dario
; APPLICANT: TARLI, Lorenzo
; APPLICANT: VITI, Francesca
; APPLICANT: BIRCHER, Manfred
; TITLE OF INVENTION: SPECIFIC BINDING MOLECULES FOR SCINTIGRAPHY, CONUGATES
; TITLE OF INVENTION: CONTAINING THEM AND THERAPEUTIC METHOD FOR TREATMENT OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: SCH-1733P2
; CURRENT APPLICATION NUMBER: US/09/512,082
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 09/300,425
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/075,338
; PRIOR FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 20
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antibody linker
US-09-512-082-20

Query Match 100.0%; Score 73; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.04; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0

QY 1 GDSSGGSGGASTG 14
|||||
DB 1 GDSSGGSGGASTG 14

RESULT 4

US-09-791-537-78345
; Sequence 78345, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 78345
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-78345

Query Match 100.0%; Score 73; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.7; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0

QY 1 GDSSGGSGGASTG 14
|||||
DB 117 GDSSGGSGGASTG 130

RESULT 5

US-09-791-537-78361
; Sequence 78361, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 78361
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-78361

Query Match 100.0%; Score 73; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.7; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0

QY 1 GDSSGGSGGASTG 14
|||||
DB 117 GDSSGGSGGASTG 130

RESULT 6

US-09-791-537-78362
; Sequence 78362, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210

;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78362
;; LENGTH: 238
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-78362

Query Match
Best Local Similarity 100.0%; Score 73; DB 22; Length 238;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGGASTG 14
Db 117 GDGSSGSGGASTG 130

RESULT 7
US-09-791-537-78364

;; Sequence 78364, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph

;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78364
;; LENGTH: 238
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-78364

Query Match
Best Local Similarity 100.0%; Score 73; DB 22; Length 238;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGGASTG 14
Db 117 GDGSSGSGGASTG 130

RESULT 8
US-09-791-537-78365

;; Sequence 78365, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78365
;; LENGTH: 238
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-78365

Query Match
Best Local Similarity 100.0%; Score 73; DB 22; Length 238;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGGASTG 14

Db 117 GDGSSGSGGASTG 130

RESULT 9
US-09-791-537-78380

;; Sequence 78380, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78380
;; LENGTH: 238
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-78380

Query Match
Best Local Similarity 100.0%; Score 73; DB 22; Length 238;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGGASTG 14
Db 117 GDGSSGSGGASTG 130

RESULT 10
US-09-791-537-78381

;; Sequence 78381, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78381
;; LENGTH: 238
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-78381

Query Match
Best Local Similarity 100.0%; Score 73; DB 22; Length 238;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDGSSGSGGASTG 14
Db 117 GDGSSGSGGASTG 130

RESULT 11
US-09-791-537-78382

;; Sequence 78382, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537

;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78382
;; LENGTH: 238
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-537-78382

Query Match 100.0%; Score 73; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDSSGSGSGASTG 14
|||||
DB 117 GDSSGSGSGASTG 130

RESULT 12
US-09-791-537-78383

;; Sequence 78383, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78383
;; LENGTH: 238
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-537-78383

Query Match 100.0%; Score 73; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDSSGSGSGASTG 14
|||||
DB 117 GDSSGSGSGASTG 130

RESULT 13
US-09-791-537-78827

;; Sequence 78827, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 78827
;; LENGTH: 238
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-537-78827

Query Match 100.0%; Score 73; DB 22; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDSSGSGSGASTG 14
|||||

DB 117 GDSSGSGSGASTG 130

RESULT 14
US-09-075-338-1
;; Sequence 1, Application US/09075338
;; GENERAL INFORMATION:
;; APPLICANT: Neri
;; APPLICANT: Tarril
;; TITLE OF INVENTION: Binding Molecule
;; FILE REFERENCE: sequence
;; CURRENT APPLICATION NUMBER: US/09/075,338
;; CURRENT FILING DATE: 1998-05-11
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 299
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-075-338-1

Query Match 100.0%; Score 73; DB 14; Length 299;
Best Local Similarity 100.0%; Pred. No. 0.88; 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

QY 1 GDSSGSGSGASTG 14
|||||
DB 123 GDSSGSGSGASTG 136

RESULT 15
US-10-276-781-1626
;; Sequence 1626, Application US/10276781
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang, et al.
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-018 (785 contig)
;; CURRENT APPLICATION NUMBER: US/10/276,781
;; CURRENT FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: 09/491,404
;; PRIOR FILING DATE: 2000-01-25
;; NUMBER OF SEQ ID NOS: 2018
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 1626
;; LENGTH: 384
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-276-781-1626

Query Match 100.0%; Score 73; DB 28; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.1; 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

QY 1 GDSSGSGSGASTG 14
|||||
DB 153 GDSSGSGSGASTG 166

Search completed: August 20, 2003, 13:13:38
Job time: 144.289 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:56 ; Search time 9.78313 Seconds

(without alignments)
137.621 million cell updates/sec

Title:	US-09-512-082-20
Perfect score:	72

Sequence: 1 GDGSSGGGGASTG 14

Scoring table: BLOSUM62

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR_76:★

```
1: plr1: *
2: plr2: *
3: plr3: *
4: plr4: *
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	58	79.5	801	2	E70824	hypothetical glyc
2	55	75.3	588	2	E70971	hypothetical glyc
3	55	75.3	1022	2	T17406	developmental glyco
4	55	75.3	1381	2	E70806	hypothetical glyco
5	55	75.3	1660	2	A70869	hypothetical glyco
6	54	74.0	183	2	PN0109	keratin-like prot
7	54	74.0	515	2	H70663	hypothetical glyco
8	54	74.0	749	2	A70812	hypothetical glyco
9	54	74.0	783	2	E70824	hypothetical glyco
10	54	74.0	1325	1	S73723	probable lipoprotei
11	54	74.0	3016	2	S77300	hypothetical prot
12	53	72.6	296	2	AA7318	gene Minc protein
13	53	72.6	296	2	IS1342	gene Minc protein
14	52	71.2	424	1	Z3BPFD	coat protein A
15	52	71.2	424	1	Z3BPM3	coat protein A
16	52	71.2	424	1	Z3BPE1	coat protein A
17	52	71.2	424	1	Z48172	variable surface
18	52	71.2	429	2	S48172	hypothetical glyco
19	52	71.2	498	2	C07070	hypothetical glyco
20	52	71.2	1079	2	B70807	hypothetical glyco
21	51	69.9	1489	2	D70807	conserved hypoche
22	51	69.9	384	2	E81689	hypothetical glyco
23	51	69.9	491	2	D70916	hypothetical glyco
24	51	69.9	542	2	T06728	pectate lyase (EC
25	51	69.9	549	2	T20720	hypothetical prot
26	51	69.9	731	2	C70974	hypothetical glyco
27	51	69.9	837	2	E70835	hypothetical glyco
28	50	69.2	1901	2	F70806	related to RNA-dir
29	50	68.3	564	2	T49322	heat-shock protein
			340	2	JN0912	

ALIGNMENTS

45	31	50	68.5	484	2	D66448
44	30	50	68.5	381	2	A35628
43	31	50	68.5	603	2	A70770
44	33	50	68.5	671	2	A35912
43	34	50	68.5	694	2	S71766
35	35	50	68.5	714	2	A70807
36	36	50	68.5	895	2	CT0089
37	37	50	68.5	1306	2	A70934
38	38	49.5	67.8	694	2	F70868
39	39	49	67.1	167	2	S21359
40	40	49	67.1	208	2	T46866
41	41	49	67.1	278	2	S39310
42	42	49	67.1	284	2	S74256
43	43	49	67.1	330	2	S74255
44	44	49	67.1	361	2	G70682
45	45	49	67.1	439	2	D70954

[illegible]

RESULT

hypocretin-like glycinergic protein Rv0747 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence.revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70824
R:Role, S.T.: Broesch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon
A:Conor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70824
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-801 <COL>
A:Cross-references: GB:AL021958; GB:AL123456; NID:93261536; PIDN:CAA17514.1; PID:9291
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0747
A:Superfamily: elastin

Query Match	79.5%;	Score 58;	DB 2;	Length 801;
Best Local Similarity	71.4%;	Pred. No. 4.9;		
Matches 10;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	GDSSSGSGGASTG	14
		1:1:1:1:1:1:1	
Db	313	GNGGNGNGGASTG	326

RESULT 2

hypodermic glycinine-rich protein RV3367 - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003
 C:Accession: F70971
 R:CoLe, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtzoyd,
 Ralndream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70971
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-588 <COL>
 A:Cross-References: GB:AL009198; GB:AL133456; NID:g3424262; PIDN:CAA15752.1; PID:e120
 A:Experimental source: strain H37RV
 A:Genetics:

A:Gene: RV3367
C:Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 75.3%; Score 55; DB 2; Length 588;
Best Local Similarity 83.3%; Pred. No. 8.7;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGAS 12
DB 429 GDGSGGAGGAS 440

RESULT 3

T17406
developmental protein - slime mold (Dictyostellium discoideum) (fragment)

C:Species: Dictyostellium discoideum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17406
R:Ratnar, N.; Loomis, W.F.
submitted to the EMBL Data Library, July 1998

A:Reference number: Z18772

A:Accession: T17406
A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-1022 <TRA>

A:Cross-references: EMBL:AF076601; NID:g3414933; PID:g3414934; PIDN:AA031540.1

C:Genetics:
A:Note: DG1105
A:Note: Intron positions not resolved (incomplete sequence)

Query Match 75.3%; Score 55; DB 2; Length 1022;
Best Local Similarity 78.6%; Pred. No. 15;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGASTG 14
DB 391 GGGSSGGSGGASG 404

RESULT 4

E70806
hypothetical glycine-rich protein RV3507 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: E70806

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: E70806
A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-1381 <COL>

A:Cross-references: GB:AL022022; GB:AL123456; NID:g3261554; PIDN:CAA17744.1; PID:g292444

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV3507

C:Superfamily: collagen alpha 1(IV) chain

Query Match 75.3%; Score 55; DB 2; Length 1381;
Best Local Similarity 64.3%; Pred. No. 19;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGASTG 14
DB 611 GDGAGGAGGAANG 624

RESULT 5

A70869

hypothetical glycine-rich protein RV2490c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: A70869

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: A70869
A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-1660 <COL>

A:Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16067.1; PID:g279

A:Experimental source: strain H37RV

C:Genetics:
A:Gene: RV2490c

C:Superfamily: collagen alpha 1(IV) chain

Query Match 75.3%; Score 55; DB 2; Length 1660;
Best Local Similarity 64.3%; Pred. No. 23;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGASTG 14
DB 1377 GDGAGGAGGAGTGTG 1390

RESULT 6

PN0109
keratin-like protein - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997

C:Accession: PN0109

R:Shivers, Y.P.; Chumakov, I.M.; Kisselev, L.L.

Mol. Biol. 24, 663-677, 1990

A:Title: Sequencing the fragment from a transcriptionally active rat genomic locus

A:Reference number: PN0109

A:Accession: PN0109
A:Molecule type: mRNA

A:Residues: 1-183 <SHV>

C:Genetics:
A:Gene: K51

C:Superfamily: Ioricin

C:Keywords: duplication

F:21-28/Region: 8-residue repeat

F:34-41/Region: 8-residue repeat

F:73-88/Region: 16-residue repeat

F:77-84/Region: 8-residue repeat

F:89-104/Region: 16-residue repeat

F:105-112/Region: 8-residue repeat

F:118-125/Region: 8-residue repeat

F:152-163/Region: 12-residue repeat

F:164-171/Region: 8-residue repeat

F:176-183/Region: 8-residue repeat

Query Match 74.0%; Score 54; DB 2; Length 183;
Best Local Similarity 71.4%; Pred. No. 3.9;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGASTG 14
DB 123 GGGSSGGGGGSSG 136

RESULT 7

H70663

hypothetical glycine-rich protein RV1840c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003
C:Accession: H70663
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98255987; PMID: 9634230
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Cross-references: GB: Z83859; GB: AL123456; NID: g3261678; PIDN: CAB06114.1; PID: g1781207
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: RV1840C
C: Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 74.0%; Score 54; DB 2; Length 515;
Best Local Similarity 64.3%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
DB 186 GDGNGAGGAGCTGTG 199

RESULT 8
A70812
hypothetical glycine-rich protein Rv0833 - Mycobacterium tuberculosis (strain H37Rv)
C: Species: Mycobacterium tuberculosis
C: Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 07-Mar-2003
C: Accession: A70812
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98255987; PMID: 9634230
A: Accession: A70812
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-749 <COL>
A: Cross-references: GB: AL022004; GB: AL123456; NID: g3261550; PIDN: CAA17639.1; PID: g291685
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: RV0833
C: Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing

Query Match 74.0%; Score 54; DB 2; Length 749;
Best Local Similarity 69.2%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDSSGGSGGAST 13
DB 204 GDGAGAGGAGCAT 216

RESULT 9
E70824
hypothetical glycine-rich protein Rv0746 - Mycobacterium tuberculosis (strain H37Rv)
C: Species: Mycobacterium tuberculosis
C: Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 07-Mar-2003
C: Accession: E70824
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID: 98255987; PMID: 9634230
A: Accession: E70824
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-783 <COL>
A: Cross-references: GB: AL021958; GB: AL123456; NID: g3261536; PIDN: CAA17513.1; PID: g291
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: RV0746
C: Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing

Query Match 74.0%; Score 54; DB 2; Length 783;
Best Local Similarity 76.9%; Pred. No. 15;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDSSGGSGGAST 13
DB 553 GDGAGGAGGAGT 565

RESULT 10
S73723
probable lipoprotein H08_orf1325 - Mycoplasma pneumoniae (strain ATCC 29342)
N: Alternate names: MG309 homolog H08_orf1325
C: Species: Mycoplasma pneumoniae
A: Variety: ATCC 29342
C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 10-Dec-1999
R: Himmelfreid, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A: Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A: Reference number: S73327; MUID: 97105885; PMID: 8948633
A: Accession: S73723
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-1325 <HIM>
A: Cross-references: EMBL: AE000038; GB: U00089; NID: g1674074; PIDN: AB96045.1; PID: g167
A: Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C: Genetics:
A: Genetic code: SGC3
C: Superfamily: Mycoplasma hypothetical protein MG309

Query Match 74.0%; Score 54; DB 1; Length 1325;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
DB 437 GGGSGGGGCTGTG 450

RESULT 11
S77300
hypothetical protein sir1403 - Synecocystis sp. (strain PCC 6803)
C: Species: Synecocystis sp.
A: Variety: PCC 6803
C: Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
R: Kaneo, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A: Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
S.
A: Reference number: S74322; MUID: 97061201; PMID: 8905231
A: Accession: S77300
A: Status: nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-3016 <KAN>
A: Cross-references: EMBL: D90907; GB: AB001339; NID: g1652618; PIDN: BAA17634.1; PID: d101
A: Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 74.0%; Score 54; DB 2; Length 3016;
Best Local Similarity 71.4%; Pred. No. 54;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
 ||| ||| ||| |||
 Db 2489 GDVSGGGGAGAG 2502

RESULT 12

A:Reference number: A93690; MUID:79136480; PMID:745987
 A:Accession: A04266
 A:Molecule type: DNA
 A:Residues: 1-424 <BEC>
 A:Cross-references: GB:V00602; GB:J02451; GB:M10731; GB:M10767; GB:M21666; GB:M21667;
 A:Experimental source: GB:V00602; strain 478, Heidelberg
 R:Goldsmith, M.E.; Konigsberg, W.H.
 A:Biochemistry 16, 2686-2694, 1977
 A:Title: Adsorption protein of the bacteriophage fd: isolation, molecular properties,
 A:Reference number: A90402; MUID:77242231; PMID:329863
 A:Accession: B04266
 A:Molecule type: protein
 A:Residues: 1-24, 'P', '26', 'P', 'P' <GOL>
 C:Comment: coat protein A is necessary for adsorption of the virion onto the F-pilus
 end of the phage particle
 C:Comment: Bacteriophages fd, M13, and f1 are male-specific filamentous coliphages.
 C:Genetics:
 A:Gene: III
 C:Superfamily: class I filamentous phage coat protein A
 C:Keywords: coat protein
 F:19-424/Product: coat protein A #status predicted <CPA>

Query Match 71.2%; Score 52; DB 1; Length 424;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
 ||| ||| ||| |||
 Db 209 GDSSGGSGGSGG 222

RESULT 13

A:Reference number: A93690; MUID:79136480; PMID:745987
 A:Accession: A04266
 A:Molecule type: DNA
 A:Residues: 1-424 <BEC>
 A:Cross-references: GB:V00602; GB:J02451; GB:M10731; GB:M10767; GB:M21666; GB:M21667;
 A:Experimental source: GB:V00602; strain 478, Heidelberg
 R:Goldsmith, M.E.; Konigsberg, W.H.
 A:Biochemistry 16, 2686-2694, 1977
 A:Title: Adsorption protein of the bacteriophage fd: isolation, molecular properties,
 A:Reference number: A90402; MUID:77242231; PMID:329863
 A:Accession: B04266
 A:Molecule type: protein
 A:Residues: 1-24, 'P', '26', 'P', 'P' <GOL>
 C:Comment: coat protein A is necessary for adsorption of the virion onto the F-pilus
 end of the phage particle
 C:Comment: Bacteriophages fd, M13, and f1 are male-specific filamentous coliphages.
 C:Genetics:
 A:Gene: III
 C:Superfamily: class I filamentous phage coat protein A
 C:Keywords: coat protein
 F:19-424/Product: coat protein A #status predicted <CPA>

Query Match 71.2%; Score 52; DB 1; Length 424;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
 ||| ||| ||| |||
 Db 209 GDSSGGSGGSGG 222

RESULT 14

A:Reference number: A93690; MUID:79136480; PMID:745987
 A:Accession: A04266
 A:Molecule type: DNA
 A:Residues: 1-424 <BEC>
 A:Cross-references: GB:V00602; GB:J02451; GB:M10731; GB:M10767; GB:M21666; GB:M21667;
 A:Experimental source: GB:V00602; strain 478, Heidelberg
 R:Goldsmith, M.E.; Konigsberg, W.H.
 A:Biochemistry 16, 2686-2694, 1977
 A:Title: Adsorption protein of the bacteriophage fd: isolation, molecular properties,
 A:Reference number: A90402; MUID:77242231; PMID:329863
 A:Accession: B04266
 A:Molecule type: protein
 A:Residues: 1-24, 'P', '26', 'P', 'P' <GOL>
 C:Comment: coat protein A is necessary for adsorption of the virion onto the F-pilus
 end of the phage particle
 C:Comment: Bacteriophages fd, M13, and f1 are male-specific filamentous coliphages.
 C:Genetics:
 A:Gene: III
 C:Superfamily: class I filamentous phage coat protein A
 C:Keywords: coat protein
 F:19-424/Product: coat protein A #status predicted <CPA>

Query Match 71.2%; Score 52; DB 1; Length 424;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
 ||| ||| ||| |||
 Db 236 GGGSGGGSGGSGG 249

Search completed: August 20, 2003, 12:42:15
 Job time : 11.7831 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: August 20, 2003, 12:13:21 ; Search time 5.31325 Seconds

(Without alignments)
123.912 Million cell updates/sec

Title: US-09-512-082-20

Sequence: 1 GDGSSGSGGAGTG 14

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSPROT_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	79.5	801	Y747_MYCTU	053810 mycobacteri
2	54	74.0	515	Y140_MYCTU	050594 mycobacteri
3	54	74.0	1325	Y309_MYCPN	P75334 mycoplasma
4	53	72.6	312	RALY_MOUSE	064012 mus musculu
5	52	71.2	424	COAA_BPRD	P03661 bacterioph
6	52	71.2	424	COAA_BPM13	P03662 bacterioph
7	52	71.2	498	Y118_MYCTU	050615 mycobacteri
8	51	69.9	1901	Y208_MYCTU	053553 mycobacteri
9	50	68.5	276	MSA2_PLAF8	059320 plasmodium
10	50	68.5	340	DJB1_HUMAN	P25685 mus sapien
11	50	68.5	481	LORI_MOUSE	P18165 mus musculu
12	50	68.5	588	T7L1_HUMAN	Q08632 mus musculu
13	50	68.5	603	YD25_MYCTU	Q10637 mycobacteri
14	50	68.5	671	HMOC_DROME	P22810 dirosophila
15	50	68.5	694	FRZ2_DROME	Q9VYX3 dirosophila
16	50	68.5	895	Z281_HUMAN	Q9Y2X9 homo sapien
17	50	68.5	2161	SHK1_HUMAN	Q9Y566 homo sapien
18	49	67.1	333	SIX3_MOUSE	Q62233 mus musculu
19	49	67.1	533	STX3_HUMAN	Q9UPE1 homo sapien
20	49	67.1	584	COAT_PAVC2	P30129 canine parv
21	49	67.1	612	UN37_CAEEL	002482 caenorhabdi
22	49	67.1	627	K2C1_MOUSE	P04104 mus musculu
23	49	67.1	722	COAT_MEVA	P27437 mink enteri
24	49	67.1	722	COAT_PAVC7	P04863 canine parv
25	49	67.1	727	COAT_FPV	P27437 mink enteri
26	49	67.1	727	COAT_FPV19	P24864 feline parv
27	49	67.1	727	COAT_PAVC19	Q11213 canine parv
28	49	67.1	737	COAT_PAVCB	P17455 canine parv
29	49	67.1	748	COAT_PAVCD	P12930 canine parv
30	49	67.1	1150	ANS1_MOUSE	P59672 mus musculu
31	48.5	66.4	463	YAE8_MYCTU	053416 mycobacteri
32	48.5	66.4	677	SP87_DICDI	P54663 dicystosell
33	48	65.8	202	TWS1_HUMAN	Q15672 homo sapien

34	48	65.8	269	1	CIT2_MOUSE	035740 mus musculu
35	48	65.8	316	1	LORI_HUMAN	P23490 homo sapien
36	48	65.8	347	1	JUNB_HUMAN	P17275 homo sapien
37	48	65.8	348	1	DJB5_MOUSE	089114 mus musculu
38	48	65.8	532	1	ZIC2_HUMAN	095409 homo sapien
39	48	65.8	593	1	K1C1_HUMAN	P13645 homo sapien
40	48	65.8	622	1	GLH1_CAEEL	P34689 caenorhabdi
41	48	65.8	763	1	Y278_MYCTU	P56877 mycobacteri
42	48	65.8	957	1	REF1_MOUSE	P48377 mus musculu
43	48	65.8	963	1	ABL_DROME	P00522 dirosophila
44	48	65.8	1520	1	SLP1_CLOTH	Q06852 clostridium
45	48	65.8	1664	1		

ALIGNMENTS

```

RESULT 1
Y747_MYCTU          STANDARD:      PRT:      801 AA.
AC 053810:
DT 30-MAY-2000 (Rel. 39, Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
DI Hypothetical PE-PGRS family protein RV0747 precursor.
GN RV0747 OR MT0772.5 OR MTV041.21.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekle A.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean A., Moulie S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fietzschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uitterback T., Feldman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL: AL021958; CAAT7514.1; -
DR EMBL: AE006968; AAK45011.1; -
DR PIR: F70824; F70824.
DR TIGR: MT0772.5; -
DR Tuberculist: RV0747; -
DR InterPro: IPR000084; PE_region.

```

DR Pfam: PF00934; PE: 1.
 KW Hypothetical protein; Repeat; Signal; Complete proteome.
 FT SIGNAL 1 30
 FT CHAIN 31 801
 FT
 FT CONFLICT 188 188 G -> S (IN REF. 2).
 FT CONFLICT 225 225 R -> G (IN REF. 2).
 FT CONFLICT 227 227 R -> G (IN REF. 2).
 FT CONFLICT 295 295 K -> R (IN REF. 2).
 FT CONFLICT 300 300 T -> G (IN REF. 2).
 FT CONFLICT 338 338 S -> I (IN REF. 2).
 FT CONFLICT 377 377 A -> P (IN REF. 2).
 FT CONFLICT 577 577 T -> A (IN REF. 2).
 FT CONFLICT 580 580 MISSING (IN REF. 2).
 SQ SEQUENCE 801 AA; 65407 MW; EA54C9BF45A0DF41 CRC64;

Query Match 79.5%; Score 58; DB 1; Length 801;
 Best Local Similarity 71.4%; Pred. No. 3.8;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
 Db 313 GNGNGNGGASTG 326
 ||:|||||

RESULT 2
 Y140_MYCTU STANDARD; PRT; 515 AA.
 ID Y140_MYCTU
 AC Q50594;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical PE-PGRS family protein Rv1840c.
 DE Rv1840c OR MT1888 OR MCTY11.04 OR MCTY355.33.
 GN Mycobacterium tuberculosis.
 OS Mycobacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares K.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RT Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleetschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RT Submitter (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z83859; CAB06114.1; .
 DR EMBL; AE007047; AAK46159.1; .
 DR PIR; H70663; H70663.
 DR TIGR; MT1888; .
 DR TubercuList; Rv1840c; .
 DR InterPro; IPR000084; PE_region.
 DR Pfam; PF00934; PE; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT SIGNAL 1 21
 FT TRANSMEM 165 185
 FT TRANSMEM 199 219
 FT SEQUENCE 515 AA; 43916 MW; 2EDC8B6C6D28B7E3 CRC64;

Query Match 74.0%; Score 54; DB 1; Length 515;
 Best Local Similarity 64.3%; Pred. No. 7.5;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
 Db 186 GDGNGGAGAGCTG 199
 |||:|||||

RESULT 3
 Y309_MYCPN STANDARD; PRT; 1325 AA.
 ID Y309_MYCPN
 AC P75334;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein MG309 homolog precursor (H08_orf1325).
 GN MPN445 OR MP397.
 GN Mycoplasma pneumoniae.
 OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 CC NCBI_TaxID=2104;
 OX
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreid R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Hermann R.,
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae."
 RT Nucleic Acids Res. 24:4420-4449(1996).
 RL -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MG307 / MG309 / MG338 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE000038; AAB96045.1; .
 DR PIR; S73723; S73723.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 KW Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 1325
 FT LIPID 28 28
 FT SEQUENCE 1325 AA; 146277 MW; 13FEF53D155ECB15 CRC64;

Query Match 74.0%; Score 54; DB 1; Length 1325;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14

DB 437 GGGSGGGGCTGTG 450

RESULT 4

RAT_MOUSE STANDARD; PRT; 312 AA.

AC 064012; 099K76; 09CXB8; 09QZK6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE RNA-binding protein Raly (hnRNP associated with lethal yellow protein)

GN (Maternally expressed hnRNP C-related protein).

OS RALY OR MERC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.

RC TISSUE-Embryo;

RX MEDLINE-93307655; PubMed-8319910;

RT Michaud E.J., Bulman S.J., Stubbbs L.J., Woychik R.P.;

RT "The embryonic lethality of homozygous lethal yellow mice (Ay/Ay) is associated with the disruption of a novel RNA-binding protein.";

RL Genes Dev. 7:1203-1213(1993).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE-94326666; PubMed-8050375;

RT Duhi D.M., Stevens M.E., Vielling H., Saxon P.J., Miller M.W., Epstein C.J., Barsh G.S.;

RT "Placental effects of the mouse lethal yellow (Ay) mutation explained by deletion of a maternally expressed gene and the simultaneous production of agouti fusion RNAs.";

RL Development 120:1695-1708(1994).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC STRAIN-C57BL/6J; TISUP-Embryonic head;

RX MEDLINE-21085660; PubMed-11217851;

RT Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H., Kuehl P., Lewis S., Matsuo Y., Nakado T., Pesole G., Quackenbush J., Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Guenrich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto S., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE-Mammary gland;

RX MEDLINE-22388257; PubMed-12477932;

RT Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Yoshiyuki S., Carninci P., Prange C.J., Rata S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Rahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.E., RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]

RP SEQUENCE OF 100-135 FROM N.A. (ISOFORM 2).

RX MEDLINE-99431566; PubMed-10500250;

RT Khredukova I., Kuklin A., Woychik R.P., Michaud E.J.;

RT "Alternative processing of the human and mouse raly genes.";

RL Biochim. Biophys. Acta 1447:107-112(1999).

CC -1- FUNCTION: Probable-RNA binding protein. Could be a heterogeneous nuclear ribonucleoprotein (hnRNP).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms-2;

CC Name=2;

CC Isoid-064012-1; Sequence-Displayed;

CC Name=1;

CC Isoid-064012-2; Sequence-VSP-005805;

CC -1- TISSUE SPECIFICITY: Widely expressed. Expressed in brain, testis, lung, spleen and kidney. Weakly expressed in liver.

CC -1- DEVELOPMENTAL STAGE: Expressed in the unfertilized egg, in the blastocyst, as well as in the developing embryo and fetus.

CC -1- DISEASE: Defects in RALY are the cause of lethal yellow mutation (Ay), a dominant allele that cause embryonic lethality when homozygous, and placental effects when heterozygous, including yellow pelage, obesity, non-insulin dependent diabetes and increased tumor susceptibility. Ay is due to a 170 kb deletion that removes all but the promoter and non-coding first exon of RALY and links them to the ASIP/Agouti gene.

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: S72641; AAC60688.1; -

DR EMBL: L17076; -; NOT_ANNOTATED_CDS.

DR EMBL: AK014356; BAB29294.1; -

DR EMBL: BC004851; AAH04851.1; -

DR EMBL: BC016587; AAH16587.1; -

DR EMBL: AF148458; AAF04488.1; -

DR MGD: MGI:97850; Raly.

DR InterPro: IPR000504; RNA_rec_mot.

DR Pfam: PF00076; rrm; 1.

DR SMART: SM00360; RRM; 1.

DR PROSITE: PS00102; RRM; 1.

DR PROSITE: PS00030; RRM_RNP_1; 1.

DR KMW Ribonucleoprotein; RNA-binding; Nuclear protein; Alternative splicing.

FT DOMAIN 21 32

FT VARSPLIC 110 125

FT FT 249 249

FT FT 281 281

FT CONFLICT 249 249

FT SEQUENCE 312 AA; 33158 MW; BFE8ED8876BFC50 CR664;

QY Query Match 72.6%; Score 53; DB 1; Length 312;

Best Local Similarity 71.4%; Pred. No. 6.1;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 GGGSSGGSGGASTG 14

Db 11 1111 1111 1
225 GDSGSGGSGGSGG 238

RESULT 5
ID COA_BPFD STANDARD: PRT: 424 AA.
AC P03661;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein A precursor (G3P).
GN IIT.
OS Bacteriophage fd.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=10864;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=478/Heidelberg;
RX MEDLINE=79136480; PubMed=745987;
RA Beck E., Sommer R., Auerwald E.A., Kurz C., Zink B., Osterburg G., Schaller H., Sugimoto K., Sugisaki H., Okamoto T., Takanami M.;
RT "Nucleotide sequence of bacteriophage fd DNA.";
RL Nucleic Acids Res. 5:4495-4503(1978).
RN [2]
RP SEQUENCE OF 19-27.
RX MEDLINE=77242231; PubMed=329663;
RA Goldsmith M.E., Konigsberg W.H.;
RT "Adsorption protein of the bacteriophage fd: isolation, molecular properties, and location in the virus.";
RL Biochemistry 16:2686-2694(1977).
RN [3]
RP STRUCTURE BY NMR OF 20-85.
RX MEDLINE=97184691; PubMed=9032075;
RA Holliger P., Riechmann L.;
RT "A conserved infection pathway for filamentous bacteriophages is suggested by the structure of the membrane penetration domain of the minor coat protein g3p from phage fd.";
RL Structure 5:265-275(1997).
RN [4]
RP ONTO THE P-PILUS OF THE HOST CELL.
RX -1 SUBUNIT: THERE ARE ABOUT FIVE COPIES OF THIS PROTEIN PER MATURE PHAGE.
CC -1 MISCELLANEOUS: THEY ARE LOCATED AT THE ADSORPTION END OF THE PHAGE PARTICLE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib.ch).
CC -----
CC EMBL: V00602; CAA23851.1; -
DR EMBL: J02451; AAA32309.1; -
DR PIR: A04266; Z3BFPD.
DR PDB: 1FGP; 16-JUN-97.
DR Pfam: PF05357; Phage.Coat.A; 2.
KW Phage recognition; Coat protein; Signal; 3D-structure.
FT SIGNAL 1 18
FT CHAIN 19 424 COAT PROTEIN A.
FT DOMAIN 19 85 N1.
FT DOMAIN 86 104 G1.
FT DOMAIN 105 235 N2.
FT DOMAIN 236 274 G2.
FT DOMAIN 235 424 CT.
FT DOMAIN 233 262 NOT ESSENTIAL FOR GENE 3 FUNCTION.
FT DOMAIN 236 274 G1Y-RICH.
FT DISULFID 25 54
FT DISULFID 64 71
FT DISULFID 206 219
FT CONFLICT 25 25 C -> P (IN REF. 2).

FT CONFLICT 27 27 A -> P (IN REF. 2).
FT HELIX 22 27
FT STRAND 31 35
FT STRAND 39 41
FT TURN 42 45
FT STRAND 46 51
FT TURN 52 53
FT STRAND 54 65
FT TURN 66 69
FT TURN 70 79
SQ SEQUENCE 424 AA: 44638 MW: 1D2FE0343AB5B2F0 CRC64;
Query Match 71.2%; Score 52; DB 1; Length 424;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GDSGSGGSGGSGG 14
Db 236 GDSGSGGSGGSGG 249
RESULT 6
ID COA_BPM13 STANDARD: PRT: 424 AA.
AC P03662;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coat protein A precursor (G3P) (Minor coat protein).
GN IIT.
OS Bacteriophage M13, and
OC Bacteriophage f1.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=10870, 10863;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Phage M13;
RX MEDLINE=81067903; PubMed=6254849;
RA van Wezenbeek P.M.G.F., Huisebos T.J.M., Schoenmakers J.G.G.;
RT "Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd.";
RL Gene 11:129-148(1980).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Phage f1;
RX MEDLINE=82211801; PubMed=6282703;
RA Beck E., Zink B.;
RT "Nucleotide sequence and genome organisation of filamentous bacteriophages f1 and fd.";
RL Gene 16:35-58(1981).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Phage f1;
RX MEDLINE=83059882; PubMed=6292494;
RA Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage f1 DNA.";
RL J. Virol. 44:32-46(1982).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.46 ANGSTROMS) OF 19-235.
RC SPECIES-Phage M13;
RX MEDLINE=98120978; PubMed=9461080;
RA Lubkowski J., Hennecke F., Plueckthun A., Wlodawer A.;
RT "The structural basis of phage display elucidated by the crystal structure of the N-terminal domains of g3p.";
RL Nat. Struct. Biol. 5:140-147(1998).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 19-105 OF COMPLEX WITH TOLA.
RX MEDLINE=99332679; PubMed=10404600;
RA Lubkowski J., Hennecke F., Plueckthun A., Wlodawer A.;
RT "Filamentous phage infection: crystal structure of g3p in complex with its coreceptor: the C-terminal domain of Tola.";
RL Structure 7:711-722(1999).
CC -1 FUNCTION: COAT PROTEIN A IS NECESSARY FOR ADSORPTION OF THE VIRION

DB	Y236	GGSGGGSGGGSEG	249
OY	1	GDSSGGSGASTG	14
DB	236	GGSGGGSGGGSEG	249
RESULT 7			
Y118_MYCTU			
ID	Y118_MYCTU	STANDARD;	PRT; 498 AA.
AC	050615;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical PE-PGSG family protein RV1818C.		
OS	RV1818C OR MT1866 OR MTCY1A1.25C.		
OS	Mycobacterium tuberculosis.		
OC	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;		
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxId=1773;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-H37RV;		
RA	MEDLINE=98295987; Pubmed=9634230;		
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,		
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,		
RA	Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,		
RA	Hornbly T., Jagels K., Krogh A., Moleen J., Moule S., Murphy L.,		
RA	Oliver S., Osborn J., Quail M.A., Rajandream M.A., Rogers J.,		
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,		
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;		
RT	"Deciphering the biology of Mycobacterium tuberculosis from the		
RT	complete genome sequence."		
RL	Nature 393:537-544 (1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CDC 1551 / Oshkosh;		
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,		
RA	Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,		
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,		
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,		
RA	Bisgal W.;		
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and		
RT	laboratory strains."		
RT	Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.		
RL	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).		
CC	-1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGSS		
CC	SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; Z78020; CAB01461.1;		
DR	EMBL; AE007045; AAK46139.1; ALT_INIT.		
DR	PIR; C70720; C70720.		
DR	TIGR; MT1866;		
DR	Tuberculist; RV1818C;		
DR	InterPro; IPR000084; PE_region.		
DR	Pfam; PF00934; PE; 1.		
KW	Prodom; PD001223; PE_region; 1.		
FT	Hypothetical protein; Transmembrane; Complete proteome.		
FT	TRANSMEM 1 21		POTENTIAL.
FT	TRANSMEM 39 59		POTENTIAL.
FT	TRANSMEM 93 113		POTENTIAL.
FT	TRANSMEM 135 155		POTENTIAL.
FT	TRANSMEM 158 178		POTENTIAL.
FT	TRANSMEM 182 202		POTENTIAL.
FT	TRANSMEM 209 229		POTENTIAL.

```

Oy      1 GDGSSGSGSGASTG 14
      ||:||||:|
Db      1831 GDAGAGGNGSGATG 1844

RESULT 9
ID      MSAR2_PLAF8      STANDARD;      PRT;      276 AA.
MSA2_PLAF8
AC      099320;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Merozoite surface antigen 2 precursor (MSA-2) (Allelic form 4).
GN      MSA2.
OS      Plasmodium falciparum (isolate 7G8).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=57266;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91218803; PubMed=2090943;
RT      Thomas A.W., Carr D.A., Carter J.M., Lyon J.A.;
      "Sequence comparison of allelic forms of the Plasmodium falciparum
      merozoite surface antigen MSA2."
      Mol. Biochem. Parasitol. 43:211-220(1990).
RL      M01. Biochem. Parasitol. 43:211-220(1990).
CC      -1- FUNCTION: MAY PLAY A ROLE IN THE MEROZOITE ATTACHMENT TO THE
      ERYTHROCYTE.
CC      -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
      (Potential).
CC      -1- DEVELOPMENTAL STAGE: DURING THE TROPHOZOITE AND SCHIZONT STAGES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
      -----
CC      EMBL, M60190; AAA29690.1; -.
DR      InterPro: IPR001136; MSA_2.1.
DR      Pfam: PF00985; MSA_2.1.
KW      Malaria; Membrane; Glycoprotein; Antigen; Signal; Repeat;
      GPI-anchor; Merozoite.
FT      SIGNAL                     1      20      POTENTIAL.
FT      CHAIN                     21      252      MEROZOITE SURFACE ANTIGEN 2.
FT      PROPEP                     253      276      HYDROPHOBIC, REMOVED DURING MATURATION
      (BY SIMILARITY).
FT      DOMAIN                     44      202      POLYMORPHIC REGION.
FT      FT      105      112      POLY-THR.
FT      CARBOHYD                     22      22      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                     36      36      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                     153      153      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                     225      225      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                     250      250      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE                     276 AA; 28172 MW; 85FA62A70400DEB2 CRC64;

Query Match      68.5%;      Score 50; DB 1; Length 276;
Best Local Similarity      64.3%;      Pred. No. 12;
Matches      9;      Conservative      2;      Mismatches      3;      Indels      0;      Gaps      0.

RESULT 10
DJB1_HUMAN
ID      DJB1_HUMAN      STANDARD;      PRT;      340 AA.
AC      P25685;
DT      01-MAY-1992 (Rel. 22, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)

```

DE DnaJ homolog subfamily B member 1 (Heat shock 40 kDa protein 1) (Heat
 shock protein 40) (Hsp40) (DnaJ protein homolog 1) (HDJ-1).
 GN DnaJb1 OR HSPF1 OR DnaJ1 OR HDJ1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=92093635; PubMed=1754405;
 RA Raabe T., Manley J.L.;
 RT "A human homologue of the Escherichia coli DnaJ heat-shock protein.";
 RL Nucleic Acids Res. 19:6645-6645(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-48.
 RC TISSUE-Placenta;
 RX MEDLINE=94071949; PubMed=8250930;
 RA Ohtsuka K.;
 RT "Cloning of a cDNA for heat-shock protein hsp40, a human homologue of
 bacterial DnaJ.";
 RL Biochem. Biophys. Res. Commun. 197:235-240(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=97131529; PubMed=8975727;
 RA Hata M., Okumura K., Seto M., Ohtsuka K.;
 RT "Genomic cloning of a human heat shock protein 40 (Hsp40) gene (HSPF1)
 and its chromosomal localization to 19p13.2.";
 RL Genomics 38:446-449(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain, and Lung;
 RX MEDLINE=922388257; PubMed=12477932;
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Uesdin T.B., Toshlyuk S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP STRUCTURE BY NMR OF 1-76.
 RX MEDLINE=96291433; PubMed=8764402;
 RA Qian Y.O., Patel D., Hartl F.-U., McColl D.J.;
 RT "Nuclear magnetic resonance solution structure of the human Hsp40
 (HDJ-1) J-domain.";
 RL J. Mol. Biol. 260:224-235(1996).
 CC -1- FUNCTION: INTERACTS WITH HSP70 AND CAN STIMULATE ITS ATPASE
 CC ACTIVITY. STIMULATES THE ASSOCIATION BETWEEN HSC70 AND H1P.
 CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

DR EMBL: X62421; CAA44287.1; -
 DR EMBL: D49547; BAA08495.1; -
 DR EMBL: D85429; BAA12819.1; -
 DR EMBL: BC002352; AA02352.1; -
 DR EMBL: BC019827; AA019827.1; -
 DR PIR: J00912; J00912.
 DR PIR: S20062; S20062.
 DR PDB: 1HDJ; 08-NOV-96.
 DR Genew; HGNC:5270; DnaJb1.
 DR MIM: 604572; -
 DR GO: GO:0003773; F:heat shock protein activity; TAS.
 DR InterPro: IPR002939; DnaJ_C.
 DR InterPro: IPR001623; DnaJ_N.
 DR InterPro: IPR003095; Hsp_DnaJ.
 DR Pfam: PF00226; DnaJ_1.
 DR Pfam: PF01556; DnaJ_C; 1.
 DR PRINTS: PR00625; DnaJPROTEIN.
 DR SMART: SM00271; DnaJ_1.
 DR PROSITE: PS00636; DnaJ_1; 1.
 DR PROSITE: PS50076; DnaJ_2; 1.
 KW Heat shock; Chaperone; 3D-structure.
 FT DOMAIN
 FT 1
 FT 70
 FT CONFLICT 11 11
 FT 13 28
 FT CONFLICT 13 28
 FT 68 68
 FT CONFLICT 81 136
 FT 150 150
 FT CONFLICT 183 183
 FT CONFLICT 320 320
 FT 6 9
 FT TURN 10 10
 FT TURN 13 14
 FT HELIX 17 29
 FT TURN 30 31
 FT TURN 33 35
 FT TURN 39 40
 FT HELIX 41 54
 FT TURN 55 56
 FT HELIX 58 66
 FT TURN 67 67
 FT HELIX 69 71
 FT SEQUENCE 340 AA; 38044 MW; 17545098B0C196DF CRC64;
 QY 1 GGGSSGGSGGASTG 14
 DB 74 GSGPSGGSGGANG 87
 Query Match 68.5%; Score 50; DB 1; Length 340;
 Best Local Similarity 64.3%; Pred. No. 15;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 RESULT 11
 LORI_MOUSE
 ID LORI_MOUSE STANDARD; PRT; 481 AA.
 AC P18165;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Loricrin.
 GN LOR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90275605; PubMed=2190691;
 RA Mehrel T., Hohl D., Kochmager J.A., Longley M.A., Bundman D.,

```

RA  Cheng C., Lichtl U., Bisher M.E., Steven A.C., Steinart P.M.,
RA  Yuspa S.H., Roop D.R.;
RT  Identification of a major keratinocyte cell envelope protein,
RT  "loricin".
RT  Cell 61:1103-1112(1990).
RN  (2)
RN  SEQUENCE FROM N.A.
RC  STRAIN-BALB/C;
RC  MEDLINE=95256248; PubMed=7738016;
RX  Diserio D., Jones A., Longley M.A., Bundman D., Rothnagel J.A.,
RA  Roop D.R.;
RT  "The proximal promoter of the mouse loririn gene contains a
RT  functional AP-1 element and directs keratinocyte-specific but not
RT  differentiation-specific expression.";
RL  J. Biol. Chem. 270:10792-10799(1995).
CC  -1- FUNCTION: MAJOR KERATINOCYTE CELL ENVELOPE PROTEIN.
CC  -1- SUBUNIT: MONOMERS ARE CROSSLINKED BY DISULFIDE AND N-(GAMMA-
CC  GLUTAMYL) LYSINE ISODIPEPTIDE BONDS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: M34398; AAA39444.1; -
CC  EMBL: U09189; AAA82152.1; -
CC  PIR: A35628; A35628.
CC  HSSP: P02876; 9WGA.
CC  MGD: MGI:96816; Lot.
CC  Keratinization.
SQ  SEQUENCE 481 AA; 37830 MW; 97349A786FE239FE CRC64;

Query Match 68.5%; Score 50; DB 1; Length 481;
Best Local Similarity 71.4%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
DB 187 GGGSSGGCGGGSGG 200

RESULT 12
T7L1_HUMAN STANDARD; PRT; 588 AA.
ID T7L1_HUMAN
AC Q9HCS4; Q9NP00;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor 7-like 1 (HMG-box transcription factor 3) (TCF-
DE 3).
DE TCF7L1 OR TCF3.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal Lung; PubMed=11085512;
RX MEDLINE=20535962;
RA Sagara N., Katoh M.;
RT "Mitomycin C resistance induced by TCF-3 overexpression in gastric
RT cancer cell line MKN28 is associated with DT-diaphorase down-
RT regulation.";
RT Cancer Res. 60:5959-5962(2000).
RN [2]
RN SEQUENCE OF 331-419 FROM N.A.
RX MEDLINE=92158676; PubMed=1741298;
RA Castrop J., van Noiren K., Clevers H.C.;
RT "A gene family of HMG-box transcription factors with homology to TCF-
RT 1.";
```

```

RL Nucleic Acids Res. 20:611-611(1992).
RN [3]
RN TISSUE-SPECIFICITY.
RP MEDLINE=99113953; PubMed=9916915;
RX Barker N., Huls G., Korinek V., Clevers H.;
RT "Restricted high level expression of Tcf-4 protein in intestinal and
RT mammary gland epithelium.";
RL Am. J. Pathol. 154:29-35(1999).
CC -1- FUNCTION: Participates in the Wnt signaling pathway. Binds to DNA
CC and acts as repressor in the absence of CTNNB1, and as activator
CC in its presence. Necessary for the terminal differentiation of
CC epidermal cells, the formation of keratohyalin granules and the
CC development of the barrier function of the epidermis (by
CC similarity). Down-regulates NOG1, leading to increased mitomycin C
CC resistance.
CC -1- SUBUNIT: Binds the armadillo repeat of CTNNB1 and forms a stable
CC complex (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Detected in hair follicles and skin
CC keratinocytes, and at lower levels in stomach epithelium.
CC -1- DOMAIN: The putative Groucho interaction domain between the N-
CC terminal CTNNB1 binding domain and the HMG-box is necessary for
CC repression of the transactivation mediated by TCF7L1 and CTNNB1
CC (by similarity).
CC -1- SIMILARITY: Belongs to the TCF/LEF family.
CC -1- SIMILARITY: Contains 1 HMG box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB031046; BAB18185.1; -
CC EMBL: X62870; CAB91064.1; -
CC HSSP: P27782; 2LEF.
CC Genew: HGNC:11640; TCF7L1.
CC MTM: 604652; -
CC GO: GO:0005634; C:nucleus; NAS.
CC GO: GO:0003700; F:transcription factor activity; NAS.
CC GO: GO:0006325; P:establishment and/or maintenance of chromatin; NAS.
CC GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC GO: GO:0030111; P:regulation of wnt receptor signaling pathway; NAS.
CC InterPro: IPR000910; HMG_12_box.
CC Pfam: PF00505; HMG_box; 1.
CC SMART: SM00398; HMG; 1.
CC PROSITE: PS50118; HMG_BOX_2; 1.
CC Transcription regulation: Activator; Repressor; Trans-acting factor;
CC Nuclear protein; DNA-binding; CTNNB1 BINDING (BY SIMILARITY).
CC DOMAIN 1 74 HMG BOX.
CC FT DNA_BIND 346 414 HMG BOX.
CC FT DOMAIN 421 427 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 5 29 GLY-RICH.
CC FT DOMAIN 117 326 PRO-RICH.
SQ SEQUENCE 588 AA; 62630 MW; 82FB0C300482A02 CRC64;

Query Match 68.5%; Score 50; DB 1; Length 588;
Best Local Similarity 64.3%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
DB 11 GGGSSGGCGGGSSAG 24

RESULT 13
YD25_MYCTU STANDARD; PRT; 603 AA.
ID YD25_MYCTU
AC Q10637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
```


Best Local Similarity 75.0%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches

1; Indels 0; Gaps 0;

QY 3 GSSGSGSGASTG 14

Db 190 GSSGSGSGSGSG 201

Search completed: August 20, 2003, 12:34:50
Job time : 7.42436 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 12:25:21 ; Search time 26.4819 Seconds

(without alignments)
136.423 Million cell updates/sec

Title: US-09-512-082-20

Perfect score: 73

Sequence: 1 GDGSSGGSGGASTG 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: SPREMBL_23:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rv1rus:*
17: sp_bacteriap:*
sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	55	75.3	537	10	Q91GJ3
2	55	75.3	588	16	O50396
3	55	75.3	628	16	Q8VJ19
4	55	75.3	928	10	O9LK43
5	55	75.3	1022	5	O76546
6	55	75.3	1381	16	O53552
7	55	75.3	1384	16	O8VJ21
8	55	75.3	1665	16	O53215
9	54	74.0	57	16	O92K72
10	54	74.0	179	10	O8LK71
11	54	74.0	184	3	O9C2X0
12	54	74.0	749	16	O53844
13	54	74.0	783	16	O53809
14	54	74.0	3016	16	P73590
15	53	72.6	362	5	O9U181
16	53	72.6	413	3	O9HEJ0

17	53	72.6	486	10	Q9AR23	Q9AR23 oryza sativ
18	53	72.6	1001	11	Q8VHK9	Q8VHK9 mus muscu
19	52	71.2	203	6	O8M1H1	O8M1H1 callitrix
20	52	71.2	203	6	O8M1B5	O8M1B5 saguinus oe
21	52	71.2	429	5	O26977	O26977 trypanosoma
22	52	71.2	509	5	O9V1W0	O9V1W0 drosophila
23	52	71.2	585	5	O9VTL5	O9VTL5 drosophila
24	52	71.2	622	16	O8VKJ6	O8VKJ6 leishmania
25	52	71.2	1016	5	O9U0V3	O9U0V3 leishmania
26	52	71.2	1079	16	O53557	O53557 mycobacteri
27	52	71.2	1217	16	O8VY9	O8VY9 mycobacteri
28	52	71.2	1472	5	O8TID7	O8TID7 dicystosell
29	52	71.2	1489	5	O33539	O33539 mycobacteri
30	51	69.9	261	4	O8N619	O8N619 homo sapien
31	51	69.9	384	16	O9PKA5	O9PKA5 chlamydia m
32	51	69.9	399	16	O98MS4	O98MS4 rhizobium l
33	51	69.9	424	11	O99N63	O99N63 mesocricetu
34	51	69.9	436	5	O9VCA8	O9VCA8 drosophila
35	51	69.9	449	5	O9V105	O9V105 drosophila
36	51	69.9	491	16	O06818	O06818 mycobacteri
37	51	69.9	501	10	O93204	O93204 arabidopsis
38	51	69.9	542	10	O9SV40	O9SV40 caenorhabdi
39	51	69.9	549	5	O19318	O19318 mus musculu
40	51	69.9	588	11	O9CSH0	O9CSH0 mus musculu
41	51	69.9	731	16	O50415	O50415 mycobacteri
42	51	69.9	775	16	O8VJ15	O8VJ15 mycobacteri
43	51	69.9	837	16	O53684	O53684 mycobacteri
44	51	69.9	1507	16	O8VJ23	O8VJ23 mycobacteri
45	51	69.9	1715	16	O8V120	O8V120 mycobacteri

ALIGNMENTS

RESULT 1

Q91GJ3 PRELIMINARY; PRT; 537 AA.
AC Q91GJ3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE P0006G03.13 protein.
GN P0006G03.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0006G03.13".
DR Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002522; BAB01613.1; .
DR Gramene; Q91GJ3; .
DR InterPro; IPR000834; zn_carboxypept.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
SQ SEQUENCE 537 AA; 58755 MW; F5069BABA60A271D CRC64;

Query Match 75.3%; Score 55; DB 10; Length 537;
Best Local Similarity 71.4%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GDGSSGGSGGASTG 14
DB 24 GDGSSGGSGGASTG 37

RESULT 2
O50396 PRELIMINARY; PRT; 588 AA.

```

AC 050396;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PGSR-family protein.
GN RV3367 OR MT004.25.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv; PubMed=9634230;
RX MEDLINE=98295967; Parkhill J., Garnier T., Churcher C., Harris D.,
RA Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekela F.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL: AL009198; CAAL5752.1; --
DR TubercuList; RV3367; --
DR InterPro: IPR002952; Eggshell.
DR InterPro: IPR000084; PE_region.
DR Pfam: PF00934; PE_1.
DR PRINTS: PR01228; EGGSHLL.
DR ProDom: PD001223; PE_region; 1.
DR PROSITE: PS00583; PFKB_KINASES_1; 1.
DR Complete proteome.
SQ SEQUENCE 588 AA; 49708 MW; 067B84097F61DAF1 CRC64;

Query Match 75.3%; Score 55; DB 16; Length 588;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGAS 12
Db 429 GDGSSGGSGGAS 440
|||||
Q8VJ19 PRELIMINARY; PRT; 628 AA.
ID 08VJ19;
AC 08VJ19;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PE_PGRS family protein.
GN MT3476.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE007154; AAK47814.1; --
DR TIGR: MT3476; --
DR InterPro: IPR002952; Eggshell.

```

```

DR InterPro: IPR000084; PE_region.
DR InterPro: IPR002173; PFKB.
DR Pfam: PF00934; PE_1.
DR PRINTS: PR01228; EGGSHLL.
DR ProDom: PD001223; PE_region; 1.
DR PROSITE: PS00583; PFKB_KINASES_1; 1.
DR SEQUENCE 628 AA; 53828 MW; 9F62B1E7CC2C567E CRC64;

Query Match 75.3%; Score 55; DB 16; Length 628;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGAS 12
Db 469 GDGSSGGSGGAS 480
|||||
Q9LK43 PRELIMINARY; PRT; 928 AA.
ID 09LK43;
AC 09LK43;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similarity to receptor protein kinase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AP000377; BAB01851.1; --
DR HSSP: P11362; IFGR.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR007090; LRR_plant.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002280; Ser_thr-kinase.
DR Pfam: PF00560; LRR_9.
DR Pfam: PF00669; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS50502; LRR_PS; 4.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; kinase; Receptor; Serine/threonine-protein kinase;
KW transferase.
SQ SEQUENCE 928 AA; 99964 MW; 13BB63639060B412 CRC64;

Query Match 75.3%; Score 55; DB 10; Length 928;
Best Local Similarity 78.6%; Pred. No. 37;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGAS 14
Db 448 GDGSSGGSGGAS 461
|||||
Q7546 PRELIMINARY; PRT; 928 AA.
ID 07546;
AC 07546;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PE_PGRS family protein.
GN MT3476.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE007154; AAK47814.1; --
DR TIGR: MT3476; --
DR InterPro: IPR002952; Eggshell.

```

```

ID 076546 PRELIMINARY; PRT; 1022 AA.
AC 076546;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Developmental protein (Fragment).
GN D61105.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Iranfar N., Loomis W.F.;
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
FT EMBL; AF076601; AAC31540.1;
FT NON_TER 1 1022
SQ SEQUENCE 1022 AA; 114463 MW; FCA8C0B8E1FD4D3 CRC64;

Query Match
Best Local Similarity 75.3%; Score 55; DB 5; Length 1022;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGSSGGSGGASTG 14
Db 391 GGGSSGGSGGASG 404

RESULT 6
O53552 PRELIMINARY; PRT; 1381 AA.
AC O53552;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PGRS-family protein.
GN RV3507 OR MTW023.14.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; AL022032; CAI17744.1;
DR Tuberculist; RV3507;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000084; PE_region.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00934; PE; 1.
DR PRINTS; PRO128; EGGSHL.
DR PROSITE; PS00583; PFKB_KINASES_1; 2.
KW Complete proteome.
SQ SEQUENCE 1381 AA; 110624 MW; CA09676BD076482 CRC64;

Query Match
Best Local Similarity 75.3%; Score 55; DB 16; Length 1381;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSSGGSGGASTG 14
Db 111 :111:111:1

```

```

Db 611 GGGAGAGAGGAANG 624

RESULT 7
O8V1Z1 PRELIMINARY; PRT; 1384 AA.
AC O8V1Z1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PE_PGRS family protein.
GN MT3612.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouiri H., Gill J., Mikula A.,
RA Bisbal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AE007163; AAK47970.1;
DR TIGR; MT3612;
DR InterPro; IPR000084; PE_region.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00934; PE; 1.
DR PROSITE; PS00583; PFKB_KINASES_1; 2.
SQ SEQUENCE 1384 AA; 110839 MW; 124AC8AAAD55EBD CRC64;

Query Match
Best Local Similarity 75.3%; Score 55; DB 16; Length 1384;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSSGGSGGASTG 14
Db 608 GGGAGAGAGGAANG 621

RESULT 8
O53215 PRELIMINARY; PRT; 1665 AA.
AC O53215;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PGRS-family (pe_PGRS family protein).
GN RV240C OR MT2564 OR MTW008.46C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).

```

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL021246; CA16067.1; ALT_INT.
 DR EMBL: AE007093; AK46868.1; -.
 DR TIGR: MT2564; -.
 DR TUBERCULIST; Rv2490c; -.
 DR InterPro: IPR000084; PE_region.
 DR InterPro: IPR002173; PfkB.
 DR Pfam: PF00934; PE_1.
 DR ProDom: PD001223; PE_region; 1.
 DR PROSITE: PS00583; PFK_KINASES_1; 1.
 DR PROSITE: PS01287; RTC; 1.
 SK Complete proteome.
 SQ SEQUENCE 1665 AA; 133700 MW; 2F30BD7766994E9 CRC64;

Query Match 75.3%; Score 55; DB 16; Length 1665;
 Best Local Similarity 64.3%; Pred. No. 67;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GDGSSGGSGGASTG 14
 Db 1382 GDGAGAGAGGTGTG 1395

RESULT 9
 092K72 PRELIMINARY; PRT; 57 AA.
 AC 092K72;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical lipoprotein transmembrane.
 GN R01981 OR SWC0435.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramepenger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591789; CAC46560.1; -.
 KW Complete proteome.
 SQ SEQUENCE 57 AA; 5230 MW; 85B5DE2FE8118302 CRC64;

Query Match 74.0%; Score 54; DB 16; Length 57;
 Best Local Similarity 71.4%; Pred. No. 2.8;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
 Db 43 GSGSGSGSGGSGG 56

RESULT 10
 08LK71

ID 08LK71 PRELIMINARY; PRT; 179 AA.
 AC 08LK71;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Extensin-like protein.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. maple arrow;
 RA Farah S., Singh J.;
 RT "A glycine max root specific promoter."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF520576; AA07351.1; -.
 DR InterPro: IPR003612; AAI.
 DR Pfam: PF00234; tryp_alpha_amyl; 1.
 DR SMART: SM00499; AAI; 1.
 SQ SEQUENCE 179 AA; 17117 MW; 7629E765ACACBEB CRC64;

Query Match 74.0%; Score 54; DB 10; Length 179;
 Best Local Similarity 76.9%; Pred. No. 9.1;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GDGSSGGSGGAST 13
 Db 82 GSGSGSGSGGSGST 94

RESULT 11
 09C2X0 PRELIMINARY; PRT; 184 AA.
 AC 09C2X0;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Hydrophobin.
 GN HCF-6.
 OS Cladosporium fulvum (Fulvia fulva).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
 OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
 OX NCBI_TaxID=5499;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Race 4;
 RX MEDLINE=21265535; PubMed=11372645;
 RA Nielsen P.S., Clark A.J., Oliver R.P., Huber M., Spanu P.D.;
 RT "HCF-6, a novel class II hydrophobin from Cladosporium fulvum."
 RL Microbiol. Res. 156:1-5(2001).
 DR EMBL: AJ251294; CAC27407.1; -.
 DR InterPro: IPR002952; EggsHeli.
 DR PRINTS: PR01228; EGGSHELL.
 DR CHAIN 17
 KW HYDROPHOBIN.
 SQ SEQUENCE 184 AA; 17147 MW; EDE7CA97356BD79E CRC64;

Query Match 74.0%; Score 54; DB 3; Length 184;
 Best Local Similarity 64.3%; Pred. No. 9.4;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDGSSGGSGGASTG 14
 Db 66 GNGGNGSGGSGGNTG 79

RESULT 12
 053844 PRELIMINARY; PRT; 749 AA.
 ID 053844;
 AC 053844;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 GN PGSS-family protein (PE_PGSS family protein).
 OS MYcobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NX NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL022004; CAA17639.1;
 DR EMBL; AE006974; AAK45096.1;
 DR TIGR; MT0854.1;
 DR Tuberculist; RV0833;
 DR Complete proteome.
 KW CONFLICT 514 514 G -> A (IN REF. 2).
 FT CONFLICT 584 584 S -> G (IN REF. 2).
 FT CONFLICT 584 584 S -> G (IN REF. 2).
 SQ SEQUENCE 749 AA; 57448 MW; C3C91307F9A8A2EC CRC64;
 Query Match 74.0%; Score 54; DB 16; Length 749;
 Best Local Similarity 69.2%; Pred. No. 40;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GDSSGSGSGAST 13
 DB 204 GDGAGGAGGAGAT 216
 RESULT 13
 ID 053809 PRELIMINARY; PRT; 783 AA.
 AC 053809;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE PGSS-family protein (PE_PGSS family protein).
 GN RV0746 OR MT0772.1 OR MT041.20.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NX NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE006968; AAK45009.1;
 DR EMBL; AE006968; AAK45009.1;
 DR TIGR; MT0772.1;
 DR Tuberculist; RV0746;
 DR InterPro; IPR000084; PE_region.
 DR InterPro; IPR000817; P10n.
 DR Pfam; PF00934; PE; 1.
 DR PRINTS; PR00341; P10N.
 DR PRODOM; PD001223; PE_region; 1.
 KW Complete proteome.
 FT CONFLICT 191 191 E -> G (IN REF. 2).
 FT CONFLICT 252 252 T -> A (IN REF. 2).
 FT CONFLICT 280 280 N -> D (IN REF. 2).
 FT CONFLICT 320 320 T -> A (IN REF. 2).
 FT CONFLICT 445 445 T -> A (IN REF. 2).
 SQ SEQUENCE 783 AA; 64158 MW; 5BEDD19039E81BFD CRC64;
 Query Match 74.0%; Score 54; DB 16; Length 783;
 Best Local Similarity 76.9%; Pred. No. 42;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GDSSGSGSGAST 13
 DB 553 GDGAGGAGGAGAT 565
 RESULT 14
 ID P73590 PRELIMINARY; PRT; 3016 AA.
 AC P73590;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein slr1403.
 GN SLR1403.
 OS Synechocystis sp. (Strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 NX NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
 Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S.,
 Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90907; BAA17634.1;
 DR InterPro; IPR003644; Calx_beta.
 DR InterPro; IPR002860; GH_BNR.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR000413; Integrin_alpha.

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE006968; AAK45009.1;
 DR EMBL; AE006968; AAK45009.1;
 DR TIGR; MT0772.1;
 DR Tuberculist; RV0746;
 DR InterPro; IPR000084; PE_region.
 DR InterPro; IPR000817; P10n.
 DR Pfam; PF00934; PE; 1.
 DR PRINTS; PR00341; P10N.
 DR PRODOM; PD001223; PE_region; 1.
 KW Complete proteome.
 FT CONFLICT 191 191 E -> G (IN REF. 2).
 FT CONFLICT 252 252 T -> A (IN REF. 2).
 FT CONFLICT 280 280 N -> D (IN REF. 2).
 FT CONFLICT 320 320 T -> A (IN REF. 2).
 FT CONFLICT 445 445 T -> A (IN REF. 2).
 SQ SEQUENCE 783 AA; 64158 MW; 5BEDD19039E81BFD CRC64;
 Query Match 74.0%; Score 54; DB 16; Length 783;
 Best Local Similarity 76.9%; Pred. No. 42;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GDSSGSGSGAST 13
 DB 553 GDGAGGAGGAGAT 565
 RESULT 14
 ID P73590 PRELIMINARY; PRT; 3016 AA.
 AC P73590;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein slr1403.
 GN SLR1403.
 OS Synechocystis sp. (Strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 NX NCBI_Taxid=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
 Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S.,
 Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90907; BAA17634.1;
 DR InterPro; IPR003644; Calx_beta.
 DR InterPro; IPR002860; GH_BNR.
 DR InterPro; IPR001343; Hemlysn_Ca_bind.
 DR InterPro; IPR000413; Integrin_alpha.

Query Match	74.0%;	Score 54;	DB 16;	Length 3016;
Best Local Similarity	71.4%;	Pred. No. 1.7e+02;		
Matches 10; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0

Qy	1	GDSSSGSGGASTG	14
Db	2489	GDGVSSGGGGAGAG	2502

RESULT 15

ID	ORG181;	PRELIMINARY;	PRT;	362 AA.
AC	Q9U181;			
AD	01-MAY-2000 (T-EMBLrel. 13, Created)			
DT	01-MAY-2000 (T-EMBLrel. 13, Last sequence update)			
D7	01-OCT-2002 (T-EMBLrel. 22, Last annotation update)			
DE	Hypothetical 37.3 kDa protein.			
GN	I3338.09.			
OS	Leishmania major.			
OC	Eukaryota; Euzoosozoa; Kinetoplastida; Trypanosomatidae; Leishmania			
OX	NCBI_Taxid=5664;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Friedlin;			
RA	Welder H., Hilbert H., Duesterhoeft A., Ivens A.C., Murphy L.,			
RA	Quell M., Rajandream M.A., Barrell B.G.;			
RL	Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Friedlin;			
RX	MEDLINE=98146435; Pubmed=9477341;			
RA	Ivens A.C., Lewis S.M., Bagherzaden A., Zhang L., Chan H.M.,			
RA	Smith D.F.;			
RT	"A physical map of the Leishmania major Friedlin genome.";			
RL	Physic Res. 8:135-145(1998).			
DR	EMBL; AL133468; CAB63136.1; -.			
KW	Hypothetical protein.			
SQ	SEQUENCE 362 AA; 37271 MW; 4FB3EC976A956C2 CRC64;			

Query Match	72.6%;	Score 53;	DB 5;	Length 362;
Best Local Similarity	71.4%;	Pred. No. 25;		
Matches	10;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0;

QY 1 GDSSGGSGGASTG 14
| | | | | : |
Db 301 GAGSSGGGGGAATG 314

Search completed: August 20, 2003, 12:40:15
Job time : 30.593 secs